keeping system must be maintained to monitor the status of the job. In addition, a simulation must be performed whose output specifies the expected time of arrival and the updated priorities for these new arrivals. The completed schedule will therefore include jobs which are to be found in machine centers and jobs which are expected during the day. Data on the priorities for all jobs and the expected time of arrival for new jobs ensure a more effective sequence decision.

The third function of this system is to prepare daily status reports which can be used to control the jobs in process. Once the job is released to the shop, its progress is reported daily. A comparison can then be made between its planned and its actual progress, and appropriate expediting can be initiated if required.

The final function that this system performs is to forecast the shop loads for each machine group. The forecast is updated weekly and shows the load for the next 10 weeks. In this way, the shop utilization can be monitored and necessary short-term changes in capacity can be planned.

Next we turn to a closer look at the process of generating a daily schedule for each machine group.

PRIORITY RULE

For every job in the central information file a priority rating is computed. To accomplish this, two files must be set up and maintained. These include:

1. Planning and Standards Library
2. Fabrication Open-order Master File

The Planning and Standards Library includes setup and run time standards for parts built in the shop. Whenever changes occur, the file can be updated. The second file contains the information shown in Fig. 15-43. The second section of the Header record is updated daily by information received through remote terminals located in the shop. Here is where the status of the job is recorded. The next section of the record, Open Operations, specifies the sequence of steps required for job completion.

From this information, the priority of the job can be calculated. As we have seen, there are several rules which can be used to determine this priority. The particular rule chosen by this shop was the *minimum slack time per operation rule*. First, a slack time per operation is computed:

$$\text{Slack time per operation} = \frac{\text{RT} - \text{ROT}}{\text{NOR}}$$

where RT = remaining calendar time to due date
R~~R~~OT = remaining operation time to order completion
NOR = number of operations remaining

OPERATIONS
MANAGEMENT

McGRAW-HILL SERIES IN MANAGEMENT

Keith Davis, *Consulting Editor*

ALLEN Management and Organization
ALLEN The Management Profession
ARGYRIS Management and Organizational Development: The Path from XA to YB
BECKETT Management Dynamics: The New Synthesis
BENNIS Changing Organizations
BENTON Supervision and Management
BERGEN AND HANEY Organizational Relations and Management Action
BLOUGH International Business: Environment and Adaptation
BOWMAN Management: Organization and Planning
BROWN Judgment in Administration
CAMPBELL, DUNNETTE, LAWLER, AND WEICK Managerial Behavior, Performance, and Effectiveness
CLELAND AND KING Management: A Systems Approach
CLELAND AND KING Systems Analysis and Project Management
CLELAND AND KING Systems, Organizations, Analysis, Management: A Book of Readings
DALE Management: Theory and Practice
DALE Readings in Management: Landmarks and New Frontiers
DAVIS Human Behavior at Work: Human Relations and Organizational Behavior
DAVIS AND BLOMSTROM Business, Society, and Environment: Social Power and Social Response
DAVIS AND SCOTT Human Relations and Organizational Behavior: Readings and Comments
DEGREENE Systems Psychology
DUNN AND RACHEL Wage and Salary Administration: Total Compensation Systems
DUNN AND STEPHENS Management of Personnel: Manpower Management and Organizational Behavior
EDMUNDS AND LETEY Environmental Administration
FIEDLER A Theory of Leadership Effectiveness
FLIPPO Principles of Personnel Management
GLUECK Business Policy: Strategy Formation and Management Action
GOLEMBIEWSKI Men, Management, and Morality
HICKS The Management of Organizations: A Systems and Human Resources Approach
JOHNSON, KAST, AND ROSENZWEIG The Theory and Management of Systems
KAST AND ROSENZWEIG Organization and Management: A Systems Approach
KOONTZ Toward a Unified Theory of Management
KOONTZ AND O'DONNELL Management: A Book of Readings
KOONTZ AND O'DONNELL Principles of Management: An Analysis of Managerial Functions
LEVIN, MCLAUGHLIN, LAMONE, AND KOTTAS Production/Operations Management: Contemporary Policy for Managing Operating Systems
LUTHANS Contemporary Readings in Organizational Behavior
LUTHANS Organizational Behavior
MCDONOUGH Information Economics and Management Systems
MCNICHOLS Policy Making and Executive Action
MAIER Problem-solving Discussions and Conferences: Leadership Methods and Skills
MARGULIES AND RAIA Organizational Development: Values, Process, and Technology
MAYER Production Management

MUNDEL A Conceptual Framework for the Management Sciences
PETIT The Moral Crisis in Management
PETROF, CARUSONE, AND MCDAVID Small Business Management: Concepts and Techniques for Improving Decisions
PIGORS AND PIGORS Case Method in Human Relations
PRASOW AND PETERS Arbitration and Collective Bargaining: Conflict Resolution in Labor Relations
READY The Administrator's Job
REDDIN Managerial Effectiveness
RICHMAN AND COPEN International Management and Economic Development
SALTONSTALL Human Relations in Administration
SARTAIN AND BAKER The Supervisor and His Job
SCHRIEBER, JOHNSON, MEIER, FISCHER, AND NEWELL Cases in Manufacturing Management
SHORE Operations Management
SHULL, DELBECQ, AND CUMMINGS Organizational Decision Making
STEINER Managerial Long-range Planning
SUTERMEISTER People and Productivity
TANNENBAUM Control in Organizations
TANNENBAUM, WESCHLER, AND MASSARIK Leadership and Organization
VANCE Industrial Administration
VANCE Management Decision Simulation

Operations Management

BARRY SHORE
Associate Professor of Management Science
Northeastern University

McGraw-Hill Book Company
New York St. Louis San Francisco Düsseldorf Johannesburg Kuala Lumpur
London Mexico Montreal New Delhi Panama Rio de Janeiro Singapore
Sydney Toronto

This book was set in Linotype 21 and Monotype 8A by The Maple Press Company.
The editors were Richard F. Dojny, Sonia Sheldon, and Edwin Hanson;
the designer was Barbara Ellwood;
and the production supervisor was Joe Campanella.
The drawings were done by Anco Technical Services.
The printer and binder was Kingsport Press, Inc.

Am

**OPERATIONS
MANAGEMENT**

Copyright © 1973 by McGraw-Hill, Inc. All rights reserved.
Printed in the United States of America. No part of this
publication may be reproduced, stored in a retrieval system,
or transmitted, in any form or by any means, electronic,
mechanical, photocopying, recording, or otherwise, without
the prior written permission of the publisher.

567890 FGFG 7987

Library of Congress Cataloging in Publication Data

Shore, Barry.
Operations management.

(Management series)
Includes bibliographies.
1. Management. 2. Operations research.
I. Title.
HD31.S46 658.4 72-10047
ISBN 0-07-057045-0

To My Mother and
the Memory of My Father

CONTENTS

Preface xv

Part One Design of the System

Chapter 1 Operational Processes and Their Associated Decisions 3
Introduction 3
Process Design, Operation, and Control 5
The Structure of Decisions 8
The Role of Management Science in Solving Structured Problems 13
Management's Use of Management Science 16
Questions 17
Problems 17
Bibliography 20

Chapter 2 Capital Investment Decisions in Process Design 21
Introduction 21
Capital Budgeting Process 22
Payback 23
Time Value of Money 25
Discounted Cash Flow Methods 26
A Systems Approach with Linear Programming 34
The Incorporation of Risk through Simulation of an Investment Proposal 42
Project Risk and the Firm's Risk 48
Summary 51
Questions 52
Problems 53
Bibliography 57

Chapter 3 Replacement Analysis 58
Introduction 58
Coterminated Plant Models 60
The Concept of Economic Life 63
Terborgh's Model 65
Terborgh's Short Cut
Benefit/Cost Ratio for Replacement Alternatives 76
Summary 84
Questions 84
Problems 85
Bibliography 88

Chapter 4 Process Design: Macro 89
Introduction 89
A Framework for Process Design 90
Case Study: The Process Design of Progressive Patient Care Facilities 93
Work Station Selection 95
Schematic Aids in Process Design 96
Work Flow: Material Handling 100
Case Study: Material Handling 100
The Make or Buy Decision 104
Case Study: Make or Buy 106
Summary 107
Questions 108
Problems 109
Bibliography 111

Chapter 5 Process Design: Micro 112
Introduction 112
An Overview of the Micro Design Process 113
Line Balancing 117
Work Methods 125
Labor Standards 130
Summary 136
Questions 136
Problems 137
Bibliography 141

Chapter 6 Human Factors in Job Design 142
Introduction 142
Engineering Design 143
Human Relations 143
Job Enlargement 144
Job Participation 148
Job Enrichment 151
Job Satisfaction, Motivation, and Job Performance 156
Summary 159
Questions 160
Bibliography 161

Chapter 7 Project Design 162
Introduction 162
Network Design 163
PERT 169
Critical Path Method 175
Scheduling with Limited Resources 181

Summary 185
Questions 186
Problems 186
Bibliography 191

Chapter 8 Facilities Layout 192
Introduction 192
Product and Process Layouts 193
Process Layout: Independence 194
Case Study: Glover Landing 208
Process Layout: Dependence 208
Summary 216
Questions 217
Problems 217
Bibliography 220

Chapter 9 Facilities Location 221
Introduction 221
Factors Which Influence Location Decisions 222
Case Study 1: Location of a Telephone Switching Exchange 225
Case Study 2: Efficiency of Ambulance Service 225
Measuring the Relative Merits of Single-facility Alternatives When the Dominant Factors Are Both Tangible and Intangible 227
Single-facility Location Decisions When the Dominant Factors Are Measurable Costs 229
Multifacility Location Decisions When the Dominant Factors Are Measurable Costs 230
Summary 241
Questions 241
Problems 242
Bibliography 245

Chapter 10 Pollution Control and Its Impact on the Process 246
Introduction 246
How Did It Get This Way? 247
Forms of Pollution 249
Why Has Waste Disposal Gotten Out of Hand? 252
Waste Disposal Costs 253
Pollution Prevention Alternatives 257
Three Strategies for Controlling Levels of Effluent Discharge 260
Case Study 1: Metro Power and Light Company 263
Case Study 2: Logan International Airport 264
Case Study 3: Good Milk Company 265
Summary 266
Questions 266
Bibliography 267

Part Two Operation and Control of the System

Chapter 11 Forecasting by Exponential Smoothing 271
Introduction 271
Major Forecasting Approaches 271
A Model for Irregular Movements: Simple Exponential Smoothing 274
Models for Trend and Irregular Movements 278
A Model for Seasonal and Irregular Movements 284
A Model for Trend, Seasonal, and Irregular Movements 287
Generality of Models 290
A Control System 291
Uncertainty of the Forecast 295
Summary 295
Questions 296
Problems 296
Bibliography 300

Chapter 12 Econometric Forecasting 301
Introduction 301
Simple Linear Regression 302
Nonlinear Relationships 311
Multivariate Regression Analysis 314
The Theory of Econometrics 319
Summary 325
Questions 326
Problems 327
Bibliography 329

Chapter 13 Aggregate Planning 330
Introduction 330
Absorbing Demand Fluctuations with Overtime, Undertime, Inventory, and Subcontracting 333
Absorbing Demand Fluctuations with Overtime, Undertime, Inventory, Subcontracting, and Backorders 336
Absorbing Demand Fluctuations with Overtime, Inventory, Hiring, and Layoff 339
Absorbing Demand Fluctuations When Overtime, Undertime, Hiring, Layoff, Inventory, and Backorder Costs Are Nonlinear 346
A Heuristic Model for Production Planning 353
Some Other Considerations in Production Planning 354
Summary 355
Questions 356
Problems 357
Bibliography 361

Chapter 14 The Management and Control of Inventory Systems 362
Introduction 362
A General Framework for Inventory Models 363

Inventory Costs 365
EOQ Model 367
Quantity Discounts 372
EOQ Model with Stockouts Allowed 375
Inventory Models under Risk 378
Aggregate Effects 385
Inventory Systems 385
Summary 386
Questions 387
Problems 387
Bibliography 391

Chapter 15 Loading, Sequencing, and Control 392
Introduction 392
A Framework for Sequencing Problems 395
n Jobs, One Machine, Flow Shop, Static Arrival Pattern, SPT 395
n Jobs, One Machine, Flow Shop, Static Arrival Pattern, Sequence-dependent Setup Times 398
n Jobs, $m = 2$ Machines, Flow Shop, Static Arrival Pattern, Minimization of Total Elapsed Time 412
n Jobs, m Machines, Flow Shop, Static Arrival Pattern, Minimization of Total Elapsed Time 414
The General n/m Job Shop Problem with a Static Arrival Pattern 416
The General n/m Job Shop Problem with Dynamic Arrival Pattern 417
A Job Shop Scheduling and Control System 419
Case Study: Court Scheduling 425
Summary 426
Questions 427
Problems 427
Bibliography 430

Chapter 16 Quality Control 432
Introduction 432
Where to Inspect 433
Sampling Risk and the Economics of Sampling 434
Acceptance Sampling by Attributes 435
Acceptance Sampling by Variables 444
Process Control 444
Case Study: Quality Control in Bank Auditing 449
Bayesian Quality Control 450
Summary 456
Questions 456
Problems 457
Bibliography 458

Chapter 17 Management Information Systems 459
Introduction 459
What Is a Management Information System and What Is It Not 460

Characteristics of Computer Systems 463
Examples of MIS 465
Integrated MIS 467
Some Management Implications 474
Questions 476
Bibliography 476

Appendix A Linear Programming and Simulation 479
Introduction 479
Mathematical Programming 480
Simulation 498
Questions 503
Problems 503

Appendix B Tables 506
Table B-1 Random Numbers 507
Table B-2 Present Value of $1.00 509
Table B-3 Present Value of Annuities: $1 per Year 510
Table B-4 Capital-recovery Factors for Interest Rates from 6 to 50 Percent 511
Table B-5 Normal Curve Areas 512
Table B-6 Summation of Terms of the Poisson Formula 513
Table B-7 Percentage Points of the *F* Distribution 516
Table B-8 Binomial Distribution 519

Index 545

PREFACE

This book is intended for use in a quarter or semester *introductory course* in operations management. With the inclusion of Appendix A it requires no background in statistics and operations research. If, however, the student does have this background, the appendix can be omitted.

The organization of the text is along functional lines and therefore approaches the subject from a traditional point of view. First the problems of systems design are considered and then the problems of operation and control. Because of this, the text is problem-oriented and not tool-oriented.

In recent years the disciplines of operations research, statistics, mathematics, and behavioral science have given some valuable insight into these problem areas. These insights, when relevant, are woven into each chapter. The objective of this strategy is to illustrate to the student the problems commonly encountered in the design, operation, and control of a complex system and the role that these management sciences can play in their resolution.

Originally the study of operational systems was limited to the factory environment. More recently, however, as our economy has shifted from a manufacturing to a service orientation and as the public sector has absorbed an increasing share of our resources, attention has broadened to include the management problems found in these areas. Although the *content* of the problems found in these diverse organizations may be quite different, we often find that the general *form* of these problems is quite similar. It is therefore quite reasonable to hypothesize that the insights both descriptive and quantitative which we have accumulated can be effectively extended to these non-factory situations. Consequently you will find that the text is not solely factory-oriented. Instead the concepts are developed without a restricted focus, and case studies found throughout the text are intended to illustrate the application of these concepts in the factory, service, educational, hospital, and government spheres.

These case studies are intended for illustration and motivational purposes, and they may be utilized by the instructor as short cases or as a springboard for other material which he might like to include in the chapter.

Sprinkled throughout the text are sections preceded and followed by a slash ornament (////). They are intended for those who wish some additional exposure to more advanced ideas. Their omission will be at little sacrifice to the continuity of the text.

The book has also been designed to encourage the use of the computer. In several chapters the student will have the opportunity to use canned computer routines to solve linear programming, integer programming, and transportation problems. There is also an opportunity for writing programs, especially for the problems which require simulation models. Countless other opportunities exist throughout the text for those who wish to integrate computer, model building, and application skills. When course time does not permit, these side trips can be omitted at no loss to the continuity of the material.

No text can be written alone. I would like to thank Professors Howard E. Thompson and Christoph Haehling von Lanzenauer for their suggestions and encouragement. For her help in typing the manuscript Sue Skalder deserves a special thank you, as do all my graduate students who labored over my rough drafts and took the time to point out mistakes and make countless suggestions in the interest of clarity and ease of reading. To my wife Charlotte and children, Tonnie and Denise, my appreciation for their patience, understanding, and encouragement.

BARRY SHORE

OPERATIONS
MANAGEMENT

part one

Design of the System

1

Operational Processes and Their Associated Decisions

INTRODUCTION

THE FACTORY SETTING: WHERE IT ALL BEGAN

In 1919 Frederick Taylor's classic work entitled *Principles of Scientific Management* was published. From that date on, the management of production processes was never to be the same again. The scientific era had indeed begun.

The major focus of Taylor's book was in the area of work standards and wage incentives. He showed that the outcome of both systematically conceived work standards and wage incentives could materially improve productivity. These work standards were set either by the careful observation of workers or by predetermined standards. Then wage incentive plans were tied to these standards in such a way that the higher-than-average worker would receive better-than-average pay.

What Taylor succeeded in doing was to link the engineering process of the plant with human nature. With a well-designed plant and a reasonable set of standards, the worker was motivated to achieve a high level of productivity. The payoff to him was higher wages, while the payoff to the firm was larger profits. Everyone could prosper under this new system.

This approach became known as *scientific management theory,* since it was based on the science of engineering and the "science" of human nature, which drew from psychology, ethics, and even physiology. Although by today's standards one might question how much science was in scientific management, one is easily convinced that Taylor's approach was quite scientific when compared with management practices which preceded him.

From a methodological point of view, of major importance was the way in which Taylor supported his hypothesis. His was not based, as were so many others, on armchair theorizing. His hypothesis was based on actual studies undertaken at Bethlehem Steel yards. He therefore employed one of the central ideas of the scientific method as we know it today: first the objective investigation of a phenomenon, then the formulation of a hypothesis, then its repeated proof by experimentation, and finally its establishment as a theory.

This successful application of the scientific approach was eventually followed by studies directed toward other operational problems. Soon a scientific body of knowledge started to emerge with application to equipment investment, replacement analysis, process design, plant layout, plant location, inventory, scheduling, and quality control. Since many of these applications required the processing of large quantities of data, it was not until the computer arrived that the full potential of this growing body of knowledge became apparent. This realization sparked an intensified level of interest from theoreticians and practitioners alike. Today the interest continues at this pace.

Not only is Taylor's scientific management approach credited with being the father of the production management school, but his experiments also served as the basis for the behavioral science school. After his study came the Hawthorne studies at Western Electric and then countless others which have subsequently contributed to a framework within which organizational, group, and individual production-oriented behavioral problems can be analyzed.

Both the scientific management and the behavioral science school have something to offer those who manage operational systems. System skills are not enough by themselves. It takes skills with both systems and human beings to effectively manage an operation, and frequently skills with human beings are the dominating requirement. The student, therefore, must acquire both. He must become aware of what each has to offer, understand how they interact, and learn their advantages and shortcomings. These then become his tools of the trade.

APPLICATIONS OUTSIDE THE FACTORY SETTING

Although the traditional focus of this field has been on the factory, the management scientist has discovered that its potential is certainly not limited to this environment. The framework can also be applied to any process whose purpose is the transformation of inputs to outputs; included would be applications directed toward a bank, airline, hospital, government, public utility, pollution control agency, automotive repair shop, supermarket, library, public education, police department, and on and on.

Among these seemingly diverse fields, there exist several problems which

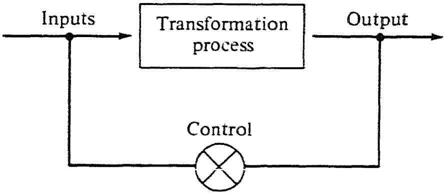


Fig. 1-1 Operational system.

are quite similar in form although quite unlike in content. For example, if we define inventory as any stock of goods which has potential economic value, then inventory theory can be applied to problems found in many fields. In a bank, cash stored for customer demand can be considered as an inventory, and seats on an airplane, seats and classrooms in a school, beds in a hospital, and switching and transmission facilities of a telephone company can all be considered inventories which are stored for future use. Our vast body of inventory theory should therefore be capable of giving some insight into the appropriate level at which these resources must be maintained.

OPERATIONS MANAGEMENT

You should now understand why the discipline originally called *production management* became known as *operations management.* The change in name merely reflected the increase in its range of applicability.

In defining operations management one should therefore be as general as possible. We might say that it is *the management of any operational system which transforms inputs into outputs.*

A very general graphic portrayal of such a system is shown in Fig. 1-1. Common to all systems are inputs, a process, outputs, and controls. Inputs usually consist of people, materials, or objects, while the process is composed of plant, equipment, and personnel.

The purpose of a system is to transform an input into an output. Generally there is some degree of repetitiveness to this procedure. That is, once the process is designed and constructed, it will be used over and over again for the same purpose.

There are two major flows in such a system. These include material and information flows. In a hospital, for example, the input would consist of the patient. He would then represent the material flow from one work center in the process to the next. Accompanying this material flow would be a substantial volume of medical records and billing information. These constitute one part of the information flow. Other information flows include data which are necessary to manage and control the system.

We now take a closer look at systems design, operation, and control.

PROCESS DESIGN, OPERATION, AND CONTROL

PROCESS DESIGN

A process can be classified as being either continuous, intermittent, or project. In a continuous process the facilities are devoted exclusively to the transformation of one type of product or service. Flow is serial and usually accomplished by conveyor. Equipment is special-purpose and has little flexibility for use elsewhere. Also prevalent is a high level of labor specialization. Typical of continuous-flow processes are automotive assembly lines and chemical and petroleum processing.

Intermittent processes contain a variety of equipment which is shared among several jobs. No job is of sufficient volume to warrant the exclusive use of one or several pieces of equipment. At the extreme no two jobs would be scheduled through the same sequence of equipment and in the same order. The equipment in these processes is usually general-purpose. The classic example of an intermittent process is a job shop.

The project process is employed whenever a sequence of operations is performed once and not repeated. Examples of this include the construction of a new building; a research and development effort designed to develop a new drug; and the design, manufacture, and deployment of a communications satellite network.

The decision as to which process orientation is best usually hinges on economic analysis. Whenever the volume is anticipated to be high and the process is expected to be long-lived, continuous processes are generally chosen. On the other hand, if the volume is expected to be low and the life of the process uncertain or short, the balance shifts to an intermittent-process orientation.

After this major choice is made, other process design decisions must follow. These include selection of equipment, detailed design of the process, and major make or buy decisions. Focus then shifts to the work that will be included at each work station. At this point human factors must be considered: should the orientation be toward job specialization, in which each person continuously performs a few routine steps, or toward job enlargement, where the worker performs more steps, has a more interesting task, and has more control over his environment?

If a project orientation is employed, the full set of jobs that comprise the project must be carefully sequenced. Of concern are those jobs which if delayed could delay the entire project. These critical jobs must be identified and carefully monitored. If some delay should occur, perhaps additional resources could be channeled in their direction.

The way in which the process facilities are laid out can certainly influence the effectiveness and efficiency of the process. A fresh look at the layout should be undertaken not only when new facilities are opened, but also whenever the process is altered, when there is a significant change in volume, or when the mix of jobs processed through the facility changes.

Facility location problems can be quite complex. For example, either the location of a thermonuclear power plant, an addition to a municipal airport, or the location of a steel mill would certainly be expected to create problems. In addition to the usual factors, such as taxes, construction costs, labor pool, land costs, and transportation costs, environmental problems represent a new dimension with which management must contend.

SCHEDULING, OPERATING, AND CONTROLLING THE PROCESS

Once the process has been designed, it must be scheduled, operated, and controlled. So that intelligent schedules can be drawn, a forecast is necessary. Sometimes the forecasting process is quite informal and relies heavily on judgment and experience. This is especially true for new products or products in a dramatically changing marketplace. In other circumstances, where the product or service has been offered for several years and historical demand data are available, statistical forecasting techniques may be employed. It is not uncommon, however, to find even these statistical forecasts tempered with good judgment.

After the forecast has been made, an aggregate plan is usually devised. This is a plan, for the next 12 months or so, which sets overall work force and production levels. Planning, however, at this stage does not entail much detail. It is a broad plan which sets the constraints within which the detailed production schedule must fit. For example, the manufacturer of a full line of TV sets would first forecast the aggregate demand for all TV sets. Then he would determine overall work force and production level strategies to meet this demand. Detail, such as the quantity of each of these models to build, when to build them, what category of labor is required, and how much of each category, represents decisions which are made at a later time when the actual production phase is closer at hand.

With thousands, hundreds of thousands, or even millions of items in stock some routine inventory planning and control system must be employed. The system must be capable of determining when forecasted orders can be met from stock, when they cannot, at what levels it is appropriate to institute replenishment orders, and how much should be ordered.

For the month in which processing is scheduled to take place, a full, detailed schedule is compiled. In manufacturing, for example, this would include the identification of all required parts, subassemblies, assemblies, and final assembly. The schedule would show when each of these is to start and when it is expected to be completed. This information is usually provided upon the job's release to the processing system.

Once the job is dispatched to the processing system, it becomes essential to monitor its progress to ensure that its completion will meet the due date. To accomplish this, elaborate computer-based data collection and processing systems are frequently employed. The job's progress is monitored by remote data input terminals located at convenient points throughout the process. In this way the most up-to-date information is assured in the control of the schedule.

Finally, the quality of the product or service must be periodically sampled to ensure that the actual output complies with process objectives. This sampling activity can take place not only at the end of the process, but also at several stages in the process. In this way, whenever a section of the process is discovered to be out of control, the appropriate corrective action can be taken before a substantial quantity of useless output is accumulated.

MANAGEMENT INFORMATION SYSTEMS

Almost every stage in the design, operation, and control of a system requires the collection, processing, and presentation of data. These data serve as an essential ingredient in the decision-making process. Without them decisions can be reached only by hunch, guess, and intuition.

When these data are part of a well-designed and integrated information system whose purpose it is to meet management's information needs, we have a *management information system*. The system should focus on the routine and occasionally on the not-so-routine decisions that fall under management's responsibility, for in these lies the most promising payoff. Above all it should be so designed that it collects, processes, and presents data which are *relevant* to these decisions. Then the decision maker can combine this source of information with any imponderables and reach a decision.

Frequently mathematical models are employed in these systems. Their function is to process all of the relevant data in a structured way and eventually find the best or good solutions to the problem.

An example of this can be seen in a job shop situation where mathematical models are used to solve production-scheduling problems. Remote input devices are often used to collect and forward data on the current status of jobs and machines in the shop at the close of each day. Then the central computer, with the help of a mathematical model, proceeds to compute a schedule. The output then serves as the foreman's schedule for the next day.

Before any management information system can be designed, it is absolutely essential that a careful study be made of the decision environment for each particular problem included as part of the information system. Good preliminary design assures that only necessary data will be collected and only useful information will be supplied to the decision maker. Accordingly, the approach that this book will take will be to analyze decision problems associated with the design, operation, and control of operational systems. Then the focus will shift to the development of a management information system.

THE STRUCTURE OF DECISIONS

COMMON CHARACTERISTICS

Regardless of whether the decision problem concerns a design, operation, or control problem, and regardless of whether the problem occurs in a factory, hospital, school, or bank, there seem to be common characteristics that extend across all decision situations. We can postulate that any such situation displays the following five characteristics:

1. Alternative strategies for employing one's resources
2. Environmental states which are out of the control of the decision maker
3. Outcomes that result from a combination of a strategy and an environmental state

4. A decision criterion which is used as the basis for comparing outcomes
5. Predictions of the likelihood that each environmental state will occur

Strategies If there were only one way in which to accomplish a task, no problem would exist. However, it is very infrequent that such is the case. In the usual situation the decision maker has to make a choice from among several strategies. For example, in the physical layout of facilities, countless different arrangements can be imagined. The decision maker must determine which of these arrangements will be most effective.

Often the alternatives are not quite so obvious, and then the decision maker must creatively generate a set. An example of this can be found in new-product development decisions. It is not obvious which products should be included in such a set.

Environmental states and outcomes The outcome of each strategy usually depends upon the prevailing environmental state. If I choose, as today's strategy, to carry an umbrella, then the outcome of this strategy will depend on whether or not it rains. If it rains, I am protected and pleased with my action. However, if it does not rain, the umbrella proves to be a nuisance. In this example the environmental states were characterized by two possible types of weather. *Environmental states are those factors which affect the outcome of a particular strategy and over which the decision maker has little or no control.* Other common environmental factors include the life of a project, competitors' strategy, market share, and economic conditions.

Frequently some knowledge exists about the likelihood of each environmental state. One might, for example, subjectively feel that the likelihood of a 10-year project life is 10 percent, of an 8-year life, 80 percent, and of a 6-year life, 10 percent. This knowledge may be based on either a subjective evaluation or historical observation.

Criterion The strategy which is finally selected is the one associated with the most attractive outcome. The degree of attractiveness, of course, depends upon the criterion by which the decision maker is comparing outcomes. It is much easier if the comparison is based on a single dimension, such as cost, profit, or sales. Whenever there are multiple objectives and the weighing of each one is unclear, it becomes quite difficult to draw comparisons.

In the management of an operational system, the criterion chosen is usually the minimization of cost. This can be defended on the grounds that the generation of sales revenue is usually out of the control of operational decision makers. Therefore, the focus of operations management is on the only variable over which control can be exercised: cost.

KNOWLEDGE OF THE DECISION ENVIRONMENT

One way of categorizing decision problems is by the state of knowledge about the environment. One of three situations can prevail: decision making under certainty, under risk, and under uncertainty. In decision making under certainty, it is assumed that only one environmental state will exist. That is, the state of the environment is known with certainty. For example, if you somehow "got the word" and knew that it was going to rain today, our problem would be considered a decision problem under certainty. As you will shortly see, decision making under certainty is not always a trivial exercise.

In decision making under risk, the probability distribution of the possible environmental states is known. For example a decision maker might have the subjective feeling that there is a 10 percent chance that the life of a project will be 5 years, an 80 percent chance that it will be 6 years, and a 10 percent chance that it will be 7 years.

Finally, in decision making under uncertainty, it is assumed that nothing is known about possible states of the environment. It is rare, however, that the decision maker is this ignorant of the future.

Now we will take a closer look at these three categories and illustrate them with a sample of the kinds of problems that will be covered in the following chapters.

Decision making under certainty Decision making under certainty occurs when there is only one possible environmental state. A decision matrix illustrating this situation is shown in Table 1-1. Here the decision maker must choose from among four alternatives. The outcomes represent the total cost for each strategy. If the objective is to minimize total cost, strategy 3 is chosen.

In this example only four strategies had to be searched to find the optimal one. Frequently the problem is much more complex: there can be a very great number of relevant strategies which should be examined.

Consider an inventory problem in which the quantity ordered at any one time can be from 1 to 1,000 pounds. With each strategy there is a cost consequence. To search each of these strategies for the best one is prohibitive. A better way does exist. We will shortly see how mathematical models can substantially improve our ability to make effective decisions under conditions such as this.

Consider another decision problem under certainty. This time the problem is to schedule 10 jobs through a chemical process. The setup time for each job depends upon its predecessor. When like products are sequenced in succes-

Table 1-1 A decision-making problem under certainty

Strategy	*State of environment*
S_1	$O_1 = 4{,}000$
S_2	$O_2 = 2{,}000$
S_3	$O_3 = 1{,}500$ ←
S_4	$O_4 = 2{,}300$

Table 1-2 Job setup times

To job *From job*										*10*
1	—	10	9	14	12	10	6	3	16	2
2	4	—	7	5	3	4	2	17	5	1
3	7	12	—	9	7	6	4	2	12	3
4	6	6	11	—	10	9	11	14	7	4
5	4	4	7	5	—	7	4	19	6	7
6	3	7	4	6	3	—	12	11	17	4
7	5	9	8	5	2	3	—	7	6	5
8	2	2	9	10	11	10	4	—	4	9
9	7	1	2	5	6	4	1	5	—	6
10	11	5	6	9	4	3	4	1	5	—

sion, the setup time is small compared with a sequence of dissimilar jobs. The setup times for a particular set of 10 jobs are shown in Table 1-2. We can see, for example, that if job 3 is followed by job 5, the setup time is 7 minutes.

One feasible sequence for processing these jobs is: 1- 3- 2- 4- 5- 6- 7- 8- 10- 9; its cost is 76. Another feasible sequence is: 1- 2- 3- 5- 4- 6- 7- 9- 8- 10 and its cost is 70. In fact, there are 10!, or somewhat over 3 million possible strategies for scheduling these 10 jobs, and each one may have a different total setup cost. As you can see, even though the state of the environment—setup time—is known with certainty, the magnitude of the problem is still overwhelming.

Still another example of decision making under certainty can be found in production scheduling. Here the problem is to absorb fluctuations in forecast demand by some combination of hiring or layoff, subcontracting, overtime, and inventories. Even if the state of the environment—demand—is assumed to be known with certainty, there remain countless strategies which combine different levels of hiring, firing, overtime, and inventories. The decision maker's objective is to uncover that strategy which minimizes production costs.

Suppose, in the last example, that demand was not known with certainty. Instead we had a good feel for the likelihood of various levels of demand. The problem would then be classified as decision making under risk. This would, in fact, manage to complicate the issue. Now we turn to a closer examination of these problems.

Decision making under risk In decision making under risk, each strategy, S_i, is associated with several outcomes, O_{ij}, one for each state of the environment, E_j. The probability that any state j will occur is P_j, and it must hold that all these probabilities sum to 1.

Consider an inventory problem where four ordering strategies are under review. One of three environmental states is likely to occur: demand for the product can be low, medium, or high. The likelihood that any of these states will exist has been estimated by management to be .1, .6, .3. The outcomes represent the consequence of a particular strategy occurring together with an environmental state. They are measured by the dollar cost of operating the inventory system under those conditions and are shown in Table 1-3. Since it is possible for stockouts to occur, the proper adjustments have been made to the costs associated with each outcome.

It is probably reasonable to expect that this kind of decision will be made over and over again, and that none of these strategies involves a substantial portion of the firm's available resources. As a consequence, whichever strategy minimizes expected costs will be selected as the best.

The expected cost of a strategy is simply the average cost that would be incurred if the process were repeated over and over again and if the likelihood of environmental outcomes was accurate. Following this rule, the expected costs (EC) for each of these strategies can be computed in the following way:

$$\begin{aligned} EC_1 &= .1(10) + .6(15) + .3(12) = 13.6 \\ EC_2 &= .1(7) + .6(19) + .3(11) = 15.4 \\ EC_3 &= .1(8) + .6(12) + .3(16) = 12.8 \\ EC_4 &= .1(12) + .6(7) + .3(20) = 11.4 \quad \leftarrow \end{aligned}$$

The selected strategy is therefore strategy 4, which has an expected cost of 11.4.

Suppose, however, that after the decision maker carefully analyzed the consequences of choosing strategy 4, he had second thoughts. If on occasion his costs should run as high as 20 (and we would expect this 30 percent of the time), he felt his boss would strongly disapprove. He therefore preferred to choose strategy 3. In this case the decision maker has an objective which is more complex than the simple minimization of expected costs. Therefore, to maximize the decision maker's *utility*, more than these costs must be considered. Frequently the objective is multidimensional, and this makes the analysis of outcomes quite complex.

Table 1-3 A decision-making problem under risk

	Environmental state		
Strategy	E_1 (*low*)	E_2 (*med.*)	E_3 (*high*)
S_1	10	15	12
S_2	7	19	11
S_3	8	12	16
S_4	12	7	20

This framework which we have developed for decision making under risk is probably appropriate for problems of the short-range to intermediate-range variety—problems where the actual occurrence of the environmental state is but a few months to a year away. As this time horizon becomes quite long, however, it becomes difficult to assess the likelihood of these environmental states.

Decision making under uncertainty If nothing is known about the likelihood with which future states of the environment will occur, the problem is classified as a decision problem under uncertainty. Under this condition, expected values for each strategy cannot be computed. Instead several other approaches exist which are often used to select a strategy. Since it is rare that decisions are made in total ignorance of the future state of affairs, and since this concept is of little help for the kinds of problems which we will be analyzing, the reader is referred elsewhere for a complete development of these ideas.[1]

THE ROLE OF MANAGEMENT SCIENCE IN SOLVING STRUCTURED PROBLEMS

The framework which we have just developed is often useful when structuring problems which decision makers face in the management of operating systems. What we need, however, beyond this ability to structure the problem is a way to solve it. Management science serves just this purpose.

FORMULATING THE MATHEMATICAL MODEL

Recall that even in the simpler case of decision problems under certainty it was shown that there could be so many strategies from which to choose that some help was needed. By formulating the decision matrix as a mathematical model we can frequently reduce the problem to a manageable size.

In this mathematical model, the outcome is called the dependent variable, while the strategy and state of the environment are independent variables. We generalize these relationships in the following way:

$$O = f(S,E)$$

The interpretation is that the outcome is a function of (depends upon) the strategy selected and the state of the environment. Suppose we make this a little more explicit by considering an actual mathematical model of an inventory problem. It is not important at this stage for you to understand how the model was generated. Later in the book you will have this opportunity.

Someone tells us that the total cost of operating an inventory system can be computed in the following way:

$$\text{TC} = DC + \frac{D}{Q}A + \frac{Q}{2}CI$$

---

[1] See David W. Miller and Martin K. Starr, *The Structure of Human Decisions*, Prentice-Hall, Inc., Englewood Cliffs, N.J., 1967.

where TC = total cost of operating the inventory system for a particular part
D = yearly demand for the part under consideration
C = per-piece cost of that part
Q = order quantity
A = fixed order costs incurred each time an order is placed. This would include typing costs, inspection, and order-expediting costs
I = cost to carry inventory in stock for 1 year as a percent of the per-piece cost

The total cost corresponds to the outcome, while the order quantity—the only variable which is totally under the control of the decision maker—corresponds to the strategy. The demand represents the state of the environment, and in this analysis we assume that demand is known with certainty. The fixed order cost, per-piece purchase cost, and inventory carrying cost are called *parameters.* They are different for each firm, depend on its economic and managerial structure, and must be carefully estimated. These parameters are necessary to make the relationship complete between the outcomes, strategies, and environmental states.

SOLVING THE MODEL

Consider a situation where the following values were determined for each of the parameters:

C = \$1
A = \$20 per order
I = 20 percent per unit

In addition, demand is estimated to be 1,000 units. Now the outcome for any strategy can be calculated. Take, for example, the strategy where $Q = 100$. The total cost can therefore be determined in the following way:

$$\mathrm{TC} = 1{,}000(1) + \frac{1{,}000}{100}(20) + \frac{100}{2}(1)(.20)$$

$$\mathrm{TC} = \$1{,}210$$

The outcome of a strategy which entails ordering 100 units at a time is \$1,210 per year. What about the outcomes of other strategies? We could compute the outcomes of many, many strategies and choose that one which has the lowest total cost. This approach, which is called *total enumeration,* would be very time-consuming and in this case very unnecessary.

Calculus By using the traditional tools of calculus, the point at which the total cost function of our inventory problem achieves its minimum can be determined. Without going into any of the extensive detail, we will present only the results. The minimum total cost Q^* will be achieved if the following quantity is ordered:

$$Q^* = \sqrt{\frac{2DA}{IC}}$$

In our example the appropriate order quantity would be

$$Q^* = \sqrt{\frac{2(1{,}000)(20)}{(.20)(1)}}$$

$$Q^* = 446$$

The total cost of this strategy is

$$\text{TC} = 1{,}000(1) + \frac{1{,}000}{446}(20) + \frac{446}{2}(1)(.20)$$

$$\text{TC} = \$1{,}049.44$$

The total cost of any other strategy would be greater.

Now it should be clear how management science can resolve this complex inventory problem. Rather than searching for the best strategy, a mathematical formula for determining Q^* can be used directly.

There are, of course, other situations where calculus is a useful means of solving a complex model. For purposes of this text, however, it has limited usefulness and consequently appears in but a few optional sections.

Mathematical programming Besides the calculus, there are other management science techniques which can be employed to resolve a variety of decision problems. Another such technique—one which will be used several times throughout the text—is mathematical programming. It is useful whenever several factors constrain the choice of strategies. In the inventory problem which we have just considered, the objective was simply to minimize total cost. There were no constraints which limited our choice of strategies. Had there been constraints, they might have limited either the space in which inventory could be placed, the funds which could be spent on inventory, or the maximum number of orders that could be placed by the purchasing department. This being the case, it would have become a problem in constrained minimization, and mathematical programming techniques could be used to find a solution.

Simulation In some problems the outcome associated with a certain strategy might be quite difficult to determine. Management science can also help us here. In a later chapter, for example, a case is presented where the decision maker is faced with a choice from among three alternatives. These include the purchase of a high-speed overhead crane, another medium-speed crane to operate alongside the present one, and retention of the status quo. The problem is to determine the outcome for each strategy.

Since the case is developed as a decision problem under risk, the determination of the outcome for each of these strategies is complex. To help us, another management science technique, known as *simulation,* is employed. By simulating the behavior of each alternative over several weeks, we can determine a measure of the performance for each strategy. These outcomes can then be compared and the best strategy chosen.

Heuristics Often simple rules of thumb, called *heuristics,* are used to solve complex problems. The solutions so obtained, however, are most likely not the best ones, but generally they are good enough. The heuristic technique is used whenever time and money do not allow for a better solution to the problem or whenever the problem is so hopelessly complex that better solutions are nearly impossible.

Perhaps the most commonly used heuristic is the "first come, first served" rule applied to many queueing situations. At a bank, for example, the customers at a teller's window are processed according to this rule. In a later chapter we will learn that, although this rule results in a quick solution to the problem, it does not generally provide the best answer. From a practical point of view, however, heuristics are often good enough.

MANAGEMENT'S USE OF MANAGEMENT SCIENCE

A characteristic of modeling is that only a limited number of factors from the real-world problem can be considered. As a result, a model of reality is created and not reality itself. In fact, it is very seldom, if ever, that humans can think in terms of reality. In most everyday situations we also model reality. For example, our views on civil disorders, campus unrest, war on poverty, pollution, politics, and supersonic aircraft are all based on models. No person has a neck long enough to stick out and thereby consider all the relevant facts, draw out all the necessary interrelationships, and finally reach a conclusion. This is beyond our capability as human beings. No, not even the President can think in terms of reality on these major issues. What we actually do is to observe the situation of interest and identify what we consider to be the key factors. These serve as the basis for our model. We then make use of the model whenever we need to analyze or process new information or to discuss the issue with others. One must always be sensitive to the fact, however, that his analysis and opinions are based on a model and not on reality. It is quite possible that some of the key factors in the environment were overlooked when the model was created. The model itself might display no internal inconsistency, but when used in the real world it may be well off the mark. All model users must therefore maintain a healthy respect for this source of research error.

In operations management the same model-building process is employed. First the situation is studied. Then the strategies, outcomes, states of nature, and parameters are identified. Next a model is constructed, the relevant data collected, and the model solved. Finally that crucial step from model solution to real-world implementation must be taken. Here the decision maker quite often must take an active role. Based on his knowledge of the model and its assumptions, he may feel that it would be inappropriate on this occasion to blindly follow the model's output. Perhaps this time one of those imponderables that was impossible to incorporate in the model is dominating the decision environment. His intervention is a necessary aspect of the decision process. It is unrealistic to expect that all of our models will be comprehensive enough to be left on their own.

QUESTIONS

Q1-1. Do the following types of organizations experience inventory problems? What kind?

(*a*) airline
(*b*) power plant
(*c*) hospital
(*d*) high school
(*e*) ski resort
(*f*) U.S. Army
(*g*) commercial bank
(*h*) Colorado Highway Department

Q1-2. Describe the content of the scheduling problems that are likely to be found in the following organizations. What criterion (or criteria) could serve to guide the final selection, and what constraints might also prove relevant?

(*a*) hospital
(*b*) automotive service center
(*c*) airline
(*d*) job shop
(*e*) TV home repair service
(*f*) urban public transportation system

PROBLEMS

P1-1. A large Midwest power company must expand one of its power plants. Of primary concern is the operating cost that will be incurred to meet future pollution standards. Since the equipment which they purchase today will last for at least 25 years, there will be little they can do when control levels are finally established. If a thermonuclear plant is installed and final pollution control standards are low, the incremental operating costs will be zero. If standards are intermediate, costs will be $10,000 per year; and if they are high, costs will be $50,000 per year. If an oil-burning plant is installed, incremental operating costs will be $3,000 for low standards, $5,000 if standards are medium, and $30,000 if standards are high. Finally, if a natural gas facility is installed, incremental costs will be $10,000 if standards are low, $11,000 if standards are medium, and $12,000 if standards are high.

(*a*) Structure the problem in terms of strategies, environmental states, and outcomes.

(*b*) If management feels that the likelihood of high, medium, and low standards is .20, .70, and .10, respectively, what is the optimal strategy?

(*c*) Do you have any reservations about the use of a cost minimization criterion?

P1-2. Sometimes the *maximin* method is used to select a strategy. Compared with expected value strategy, it is a cautious and pessimistic way in which to make a selection. The method employed is to identify the worst possible outcome associated with each strategy. Finally the best of this lot is chosen.

(*a*) Given the following table of profits by strategy and state N, solve for the maximin strategy.

Strategy	N_1	N_2	N_3
S_1	−2	3	5
S_2	7	2	6
S_3	5	6	4
S_4	5	2	1

(*b*) If the likelihood of occurrence for states N_1, N_2, and N_3 is .20, .30, and .50, respectively, what choice will maximize the expected monetary profit?

P1-3. On occasion the *maximax* method is used to select a strategy. It represents an optimistic approach to strategy selection. The method used to arrive at this choice is to identify the best possible outcome for each strategy. Finally, the best of this set is chosen.

(*a*) For the problem given in Prob. P1-2, solve for the maximax strategy.

(*b*) How does this solution compare with the maximin and expected monetary profit solutions?

(*c*) If you were going to encounter this problem only once, and if your resources were very limited, which might you choose?

P1-4. The Ace Trucking Company must route one of its trucks through five major cities once every week. The driving times are given in the following table.

From \ *To*	*A*	*B*		*D*	*E*
A			4		
B			5		
C			—		
D			4		
E			7		

(*a*) What constitutes an alternative?

(*b*) Is this a decision problem under certainty, risk, or uncertainty?

(*c*) How many alternative strategies are there?

(*d*) Is there an implied criterion in the statement of the problem? What is it? Might there be other criteria which could guide the final choice?

P1-5. The payoff for each given level of a particular investment will depend upon the market share achieved. The likelihood of a 10, 15, 25, and 35 percent market share is .20, .30, .40, and .10, respectively. Given the following table of outcomes, which strategy should be selected in order to maximize expected payoff?

Investment level	*Market share*			
	N_1	N_2	N_3	N_4
$S_1 = 1{,}000$	200	400	600	800
$S_2 = 2{,}000$	100	300	500	600
$S_3 = 3{,}000$	−300	200	800	1,000
$S_4 = 4{,}000$	−100	100	1,000	2,000

P1-6. Chris Houvras has spent 175 straight days at the track waiting for his favorite horse to run. Today is the day. Just before post time he rereads *The Turf* carefully and estimates that his horse has a 70 percent chance of winning, a 20 percent chance of placing, a 5 percent chance of showing, and a 5 percent chance of losing. The table below shows the consequence of placing a \$2 win, place, or show bet.

Bet	**Result**			
	Win	*Place*	*Show*	*Lose*
Win	12	−2	−2	−2
Place	8	8	−2	−2
Show	4	4	4	−2

(*a*) Based on the expected profit criterion, which strategy should he choose?

(*b*) Have any strategies been left out?

(*c*) What conclusion would have been reached had the maximin method been used? Do you think it likely that gamblers would act in this way?

(*d*) Suppose this horse runs every day for several weeks. Further assume that all the numbers do not change. What should Chris be willing to pay for "perfect information" before the race starts?

P1-7. Three jobs, 1, 2, and 3, must be processed on two machines. Each job must pass through first A and then B. Process times for each job are given in the following table.

Job	*Machine A*	*Machine B*
1	2	6
2	4	2
3	3	4

(*a*) What are the alternative schedules?

(*b*) Which one minimizes the elapsed time?

(*c*) Can you suggest criteria other than elapsed time that could guide the choice of a strategy?

P1-8. The Pure Hogary Hog Farm purchases its feed monthly. It can choose any combination of mixes A and B that meets the nutritional requirement of the hogs. Usually the price for each mix varies from month to month. This month mix A sells for 12 cents per pound while mix B is 10 cents per pound. The following table gives the nutritional content of each mix and the minimum requirements for the four-week planning period.

	Mix A	*Mix B*	*Minimum requirement*
Vitamin no. 1	10 units/no.	4 units/no.	1,000 units/month
Vitamin no. 2	12	6	1,200
Minerals	4	8	800

(*a*) What constitutes the alternatives?
(*b*) What are the constraints?
(*c*) Is the criterion a reasonable one?
(*d*) How many pounds of mix A and mix B should be purchased this month?

BIBLIOGRAPHY

Buffa, Elwood S.: *Modern Production Management*, John Wiley & Sons, Inc., New York, 1969.

Darden, John, F. Warren McFarlan, and William M. Zani: *Managing Computer-Based Information Systems*, Richard D. Irwin, Inc., Homewood, Ill., 1971.

Miller, David W., and Martin K. Starr: *Executive Decisions and Operations Research*, Prentice-Hall, Inc., Englewood Cliffs, N.J., 1969.

Taylor, F. W.: *Principles of Scientific Management*, Harper & Brothers, New York, 1919; reprinted in *Scientific Management*, Harper & Brothers, New York, 1947.

Timms, Howard L., and M. F. Pohlen: *The Production Function in Business*, Richard D. Irwin, Inc., Homewood, Ill., 1970, chap. 5.

Wagner, Harvey M.: *Principles of Operations Research*, Prentice-Hall, Inc., Englewood Cliffs, N.J., 1969.

2

Capital Investment Decisions in Process Design

INTRODUCTION

Many decision situations which occur during the process design stage require the evaluation and comparison of capital expenditure proposals. The purchase of machinery for a new assembly line is one example of this. It is therefore essential that the basic concepts of investment analysis be understood before we delve into the topic of process design.

Investment decisions have traditionally been at the center of attention. There are, of course, good reasons for this. First, it is the allocation of capital that starts the process which eventually leads to the commitment of labor and materials. Second, once capital has been allocated, the investment is irreversible; and third, the usually large magnitude of the capital outlay makes the investment decision a major influence on the future profitability of the firm. In fact, it is the very mix of these capital investments and their relative riskiness that plays a large role in establishing the value that stockholders place on the stock. Therefore a firm with a rather conservative mix of investments and with relatively slow growth prospects might receive a low price-to-earnings multiple in the stock market. It appears, then, that the concerns in capital management are to maintain an adequate rate of return on investment so that the future profitability of the firm is assured and to maintain the current policy regarding the risk complexion of the present set of investments.

In this chapter, we will look at the way in which the productivity of investments—as measured by the payback, benefit/cost ratio, and internal rate of return—is determined, then at a systems approach for selecting a set of investments in the face of capital budget and manpower constraints, and finally at the way in which risk can be considered.

CAPITAL BUDGETING PROCESS

Consider the decision problem faced by the Getwell Electronics Company. It is a firm which manufactures electronic patient-monitoring equipment used by hospitals in intensive care wards and in operating rooms. Recently several new-product suggestions have been made to the president, Mr. Pacemaker. These products he feels would be complementary to those now in the product line, and at first glance they all look promising. Since they all require the purchasing of new production machinery, a substantial investment would have to be made in each case. The benefits accruing from these investments, on the other hand, include streams of product revenues extending well into the future. Each of these products, therefore, has both benefits and costs.

The capital that Mr. Pacemaker has allotted to these new investments is limited. It will therefore be impossible to accept all of them this year. Selection must therefore take place in such a way that efficient use is made of these scarce funds, and Mr. Pacemaker decides to set up a meeting with his finance department to discuss the problem.

At the meeting Mr. Windfall, the vice president of finance, proves to be an excellent source of help. He feels that the projects should be ranked according to their profitability, and suggests using the benefit cost method to accomplish this. First the total costs and revenue streams must be estimated. With this advice Mr. Pacemaker proceeds to collect the relevant data.

In this parable we can see the three major steps in the capital-budgeting process:

1. The generation of investment alternatives
2. The estimation of benefits and costs
3. The selection of the alternative

Although there is no magic formula for the generation of investment alternatives, it is necessary for management to make a conscious effort to channel effort in this direction. For without a substantial set of investment alternatives from which to select, the long-run profitability of the firm is seriously jeopardized.

A very difficult but nonetheless necessary task is to estimate the benefits and costs associated with each investment. Benefits usually accrue over several time periods, and it may therefore be necessary to consider the uncertainty of the future in our analysis. This, however, makes things rather complicated, and this aspect of decision making under risk will be considered later in the chapter. First we will assume that future benefits are known with certainty. Although in the real world this seldom occurs, some useful insights and many times adequate answers can be found in this way.

The final step is to select the alternative, and it is on this step that we now turn our focus. The selection of the appropriate alternative depends, of course, on the criterion chosen. We assume here that the objective of the firm is to maximize the value of the firm to its shareholders, and this value can be measured by the average share price of the stock. The average is used, since day-to-day fluctuations are not considered important, and the interest is in the long run rather than the short run. Decisions, of course, can be made to benefit the short run at the expense of the long run. For example, research and development expenditures could be curtailed so that a short-run profit increase could be shown. The long-run implications of such a policy, however, may be disastrous. Used properly, the discounted cash flow methods will achieve this objective, and we will shortly turn our attention to them. First a simpler and widely used method for selecting alternatives will be considered.

PAYBACK

Perhaps the simplest and most popular method used by industrial practitioners is the *payback* method. The payback is simply the number of years it takes to recover the initial investment. Consider one of the investments being considered by Mr. Pacemaker of the Getwell Electronics Company, an electronic cardiograph machine used to record the patient's heartbeat for diagnosis. It is estimated that the product will generate revenues for at least 4 years. Since this is in a field of rapid technological change, any forecast beyond the 4-year period would be considered of questionable value. Profit contribution, Zn, in year n is defined as the revenue in that year less expenses. The estimates for this product are

n	Zn
1	\$100(000)
2	200
3	200
4	100

If the project costs \$400(000), the payback period is 2½ years. That is, it takes 2½ years to recover the cost.

The decision to accept or reject the proposal depends upon how the computed payback figure compares with a standard. For example, if the payback standard were 3 years, the proposal would be accepted; with a computed payback of 2½ years, it takes even less time to recoup the investment than the standard. Therefore, the decision rule is: accept if the computed payback is less than the standard; otherwise, reject it.

There are, however, two serious shortcomings in the payback method. Both are best illustrated by examples. First consider that we have a choice between investments A and B.

Investment A		*Investment B*	
n	*Zn*	*n*	*Zn*
1	50	1	50
2	100	2	100
3	200	3	200
		4	100
		5	50
		6	50

Both investments cost $150 each; therefore the payback period for each is 2 years. If the payback standard is 4 years, the payback criterion tells us that we should be indifferent between the two. It is clear that we should not be indifferent. Although the payback period for both is identical, investment B has cash flows that extend well beyond the payback period. Therefore the first shortcoming of the payback method is that no consideration is taken of flows past the payback year.

The second problem can be seen from another example. Consider two investments, A and B. The flows from these investments are shown in Table 2-1. If we assume that both investments cost $450, the payback will be the same for both of them, 3 years. Are we, in fact, indifferent to both investments? We should not be, because the bulk of investment B is recovered earlier than the bulk of investment A. And since money now is more valuable than an equivalent sum later, investment B would be a better choice. Why is money

Table 2-1 Payback example

Investment A		*Investment B*	
n	*Zn*	*n*	*Zn*
1	50	1	300
2	100	2	100
3	300	3	50
4	100	4	100
5	100	5	100
6	100	6	100

now more valuable than an equivalent sum later? Because a dollar can be placed in a savings bank and at the end of a year be worth \$1.05 if the interest rate is 5 percent. Money, then, has time value. This is the second shortcoming of the payback method: it does not consider the time value of money.

Before jumping into the techniques that do take into account the time value of money, perhaps a few good words might be said about the payback method. First, it is simple to understand and simple to use. Secondly, it encourages projects with short payback; therefore the risk associated with accepting projects which take a long time in paying off is reduced. Finally, a firm that is pressed for cash might use this as a method to ensure rapid cash recovery. There are better methods to handle these problems, and we will see some of them shortly. First things first; let us overcome the problem associated with the time value of money.

TIME VALUE OF MONEY

It is characteristic of most investments that benefits will flow over several years. Since a dollar today is not the same as a dollar received one year from now, some adjustment must be made to these flows.

FUTURE VALUE OF A PRESENT AMOUNT

To illustrate the time value of money, consider a dollar placed in a savings account at the beginning of a year. At the end of the year, with the interest rate at 5 percent, the value of the initial sum is

$$P_1 = 1 + .05(1) = 1.05$$

where P_1 represents the value of the sum at the end of the first year. In general terms, this can be written as

$$P_1 = P_0 + iP_0$$

or

$$P_1 = P_0(1 + i) \tag{2-1}$$

where P_0 = initial sum
i = interest rate

If the process continues and the \$1.05 is put away for another year,

$$P_2 = 1.05 + .05(1.05) = 1.10$$

In general notation,

$$P_2 = P_1 + iP_1$$

or

$$P_2 = P_1(1 + i) \tag{2-2}$$

Since $P_1 = P_0(1 + i)$, we can write Eq. (2-2) as

$$P_2 = P_0(1 + i)(1 + i)$$

$$P_2 = P_0(1 + i)^2$$

In fact, the future value in any year n can be written as

$$P_n = P_0(1 + i)^n \qquad (2\text{-}3)$$

PRESENT VALUE OF A FUTURE AMOUNT

We have just determined the future value of a present amount. For our purposes the more relevant question is the present value of a sum received in the future. For example we might be interested in the present value of a dollar to be received in 1 year. Using the notation of Eq. (2-1) we are interested in P_0, the present value of a sum to be received in the future, P_1. Therefore the formula for present value can be written as

$$P_0 = P_1 \cdot (1 + i)^1$$

As an example, suppose that the interest rate is 10 percent. Then the present value of a dollar received at the end of 1 year will be

$$P_0 = 1\left[\frac{1}{(1. + .10)^1}\right] = 1\left(\frac{1}{1.1}\right) = 1\ (.909) = .909$$

In general terms, the present value for a sum received in any year, n, can be written as

$$P_0 = P_n\ (1 + i)^n$$

The ratio $1/(1 + i)^n$ is called the *discount factor,* and values for it can be read from the discount table (Table B-2) found in Appendix B.

Example: Find the present value of \$150 received at the end of 5 years if the interest rate is 8 percent.

The discount factor for 5 years at 8 percent can be read from the table as .681.

$$P_0 = P_i\ \frac{1}{(1 + .08)^5}$$

$$P_0 = 150(.681)$$

$$P_0 = 102.15$$

DISCOUNTED CASH FLOW METHODS

PRESENT VALUE METHOD

There are two variants to the discounted cash flow method: the present value and the internal rate of return methods. In the present value method, future flows are discounted to the present and then compared with the cost of the investment. The interest rate used to discount this flow is the firm's cost of acquiring funds: more will be said about this later in the chapter.

As an example of this method, consider a project with the following stream of benefits extending from year 1 through year 4.

	Z_n
1	50
2	100
3	150
4	200

If the cost of obtaining funds is 10 percent, the relevant discount factors can be read from the following table.

n	(1) Z_n	(2) $\frac{1}{(1+i)^n}$	(1) × (2)
1	50	.909	45.45
2	100	.826	82.60
3	150	.751	112.65
4	200	.683	136.60
			PV = 377.30

Next the present value of each of the flows is determined, (1) × (2), and finally the present value of the project PV is determined by summation.

Before we make any decisions about this investment, let us write out this present value method, using general notation. At this point it is necessary to introduce sigma (Σ) notation. The sign Σ means "sum what follows." For example, $\sum_{n=1}^{4} Z_n$ means that Z_n is to be summed from $n = 1$ to $n = 4$. Therefore the figure at the bottom of the summation sign tells us where to begin, while the figure on top tells us where to end. Writing this out, we have

$$\sum_{n=1}^{4} Z_n = Z_1 + Z_2 + Z_3 + Z_4$$

In our example

$$\sum_{n=1}^{4} Z_n = 50 + 100 + 150 + 200 = 500$$

To get the present value of the flow we did the following:

$$\mathrm{PV} = \sum_{n=1}^{4} Z_n \frac{1}{(1+i)^n}$$

$$\mathrm{PV} = Z_1 \frac{1}{(1+i)^1} + Z_2 \frac{1}{(1+i)^2} + Z_3 \frac{1}{(1+i)^3} + Z_4 \frac{1}{(1+i)^4}$$

Therefore the general form of the present value formula for a series of flows up to year N is:

$$\mathrm{PV} = \sum_{n=1}^{N} Z_n \, (1+i)^n$$

If there is plenty of cash around, then we accept all projects whose present value exceeds their cost. In our example the cost of the project is \$300. Therefore the present value of the benefits is greater than the cost incurred. The decision is to accept.

Another way of stating this decision rule is to define net present value as present value less the cost C of the project.

$$\mathrm{NPV} = \sum_{n=1}^{N} Z_n \frac{1}{(1+i)^n} - C$$

and if the NPV is greater than zero, accept the investment. In our example NPV = 377.30 — 300.00 = 77.30.

INTERNAL RATE OF RETURN METHOD

The second variant of the discounted cash flow method is the internal rate of return. The internal rate of return r is a very special interest rate. It is that interest rate which equates the present value of the benefit stream to the cost of the investment. Returning to our example, we found that an interest rate of 10 percent reduced the benefit stream to 377.30. Since this is in excess of our cost, we must *increase* our discount rate so as to reduce the present value of the stream. There is no way to determine the "right" interest rate beforehand. Only the frustrating method of trial and error will give us the answer. So let us try 16 percent.

		16%		*18%*		*20%*	
n	Z_n	$\frac{1}{(1+i)^n}$	$Z_n \frac{1}{(1+i)^n}$	$\frac{1}{(1+i)^n}$	$Z_n \frac{1}{(1+i)^n}$	$\frac{1}{(1+i)^n}$	$Z_n \frac{1}{(1+i)^n}$
1	50	.862	43.10	.847	42.35	.833	41.65
2	100	.743	74.30	.718	71.80	.694	69.40
3	150	.641	96.15	.609	91.35	.579	86.85
4	200	.552	110.40	.516	103.20	.482	96.40
			323.95		308.70		294.30

Still not large enough. Let's try 18 percent. Again we are not exactly at $300 but very close. Try 20 percent. Now we are under $300 at $294.30. Therefore we conclude that the internal rate of return of this project is approximately 19 percent.

If we assume once again that there is no capital-budget constraint, our decision rule will be to accept all projects whose internal rate of return is greater than the cost of acquiring funds. In this case the internal rate of return at 19 percent is greater than the cost of additional funds at 10 percent. Therefore the internal rate of return method agrees with the conclusion of the present value technique: accept.

METHODS OF RANKING INVESTMENTS

The decision rules just developed apply only if there is unlimited capital. Since this is usually not the case, it is necessary to rank investment alternatives in order of profitability and proceed down the list, selecting investments until the budget is exhausted.

Using the present value method, investments are ranked as the ratio R of benefits to costs C, or

$$R = \frac{\sum_{n=1}^{N} Z_n[1/(1+i)^n]}{C}$$

If the ratio is greater than 1, the project is profitable, and of course, the larger the ratio, the better. In our example the benefit/cost ratio is 377.30/300.00 = 1.24. Given a set of investment alternatives, the investments are ranked in order of their benefit/cost ratios, and projects are selected from this list until the capital budget is exhausted.

In the internal rate of return method, the projects are simply ranked by their internal rates of return r. Projects with the highest internal rates of return are selected first, and selection from the list is continued until the budget is exhausted.

//// COST OF CAPITAL

In theory, the firm's cost of acquiring funds, or its cost of capital, is that rate of return on a project that will not change the market price of the firm's stock.[1] In the selection of an investment it therefore becomes the minimum rate of return allowed on a prospective investment. A rate larger than this will increase the market price of the stock while a rate lower will reduce it. The importance, then, of determining the cost of capital for the firm cannot be overemphasized.

A way in which the cost of capital is measured is to take a weighted average of cost of all the sources of funds. This places a cost on the pool of money from which investments are financed.

---

[1] James C. Van Horne, *Financial Management and Policy*, Prentice-Hall, Inc., Englewood Cliffs, N.J., 1969, p. 110.

These funds are generated from at least three sources. They include (1) debt, (2) equity, and (3) retained earnings. The cost of each of these sources is that discount rate which equates the value of the funds received to the present value of the expected outflows.

The cost of debt is easiest to determine and can be generalized as that rate k_d that will discount the future contractual obligations of the bond to equal the proceeds of the issue.

$$P_d = \frac{I_1}{(1+k_d)} + \frac{I_2}{(1+k_d)^2} + \cdots + \frac{I_n}{(1+k_d)^n} + \frac{R_n}{(1+k_d)^n}$$

where P_d = funds received from debt issue
I_n = interest payment in period n
R_n = repayment of capital in period n

In the case of equity capital, the dividends represent payment for the use of funds. The cost can be determined as the rate k_e which will equate the present market price of the stock with the future stream of expected dividends plus the market price of the stock when the investor sells it at the end of period n.

$$P_e = \frac{D_1}{(1+k_e)} + \frac{D_2}{(1+k_e)^2} + \cdots + \frac{D_n}{(1+k_e)^n} + \frac{P_n}{(1+k_e)^n}$$

where P_e = present market price of stock
D_n = dividends in period n
P_n = market price at end of period n

The cost of retained earnings represents the dividends that are forgone by common stockholders. Therefore, at a minimum, the cost of retained earnings is the cost of equity capital.

Now a weighted cost of capital can be determined. Suppose that the firm's capital structure looked like the following:

	Amount	*Proportion*
Debt	\$ 25 million	25%
Common stock	25	25
Retained earnings	50	50
	\$100 million	100%

Furthermore, the costs for these methods of financ

	Cost (%)
Debt (k_d)	4.5
Common stock (k_e)	10.0
Retained earnings (k_e)	10.0

The weighted cost of capital can then be determined in the following way.

	(1) *Fraction of total*	(2) *Cost (%)*	(1) × (2) *Weighted cost (%)*
Debt	0.25	4.5	1.1
Common stock	0.25	10.0	2.5
Retained earnings	0.50	10.0	5.0
Cost of capital			8.6%

The practical problems involved in determining the cost of capital can be overwhelming. Consider the problem of estimating the expected dividend stream in the cost of equity capital, or the problem of whether the book value or the market value should be used in determining the proportions. These and other problems are beyond the scope of this book. For our purposes, we will assume that the hard work has already been undertaken and that we have in our possession the cost of capital k. ////

CONSISTENCY BETWEEN DISCOUNTED CASH FLOW METHODS

If the outcomes of the present value and internal rate of return methods agreed, which one was used would simply be a matter of taste. Unfortunately this is not the case.

When a single investment is being evaluated, the present value and internal rate of return methods *generally* lead to the same conclusion. This can be seen by looking at the net present value profile. The net present value profile is a plot illustrating the way in which the NPV of the investment varies with changes in the interest rate. The net present value profile for our example is shown in Fig. 2-1. At a zero rate of discount the net present value is simply the undiscounted sum of the cash flows less the cost of the project. The remainder of the points represent the net present values at 10, 16, 18, and 20 percent. The curve drawn through these points is the net present value profile. Since the internal rate of return is that interest rate which reduces the

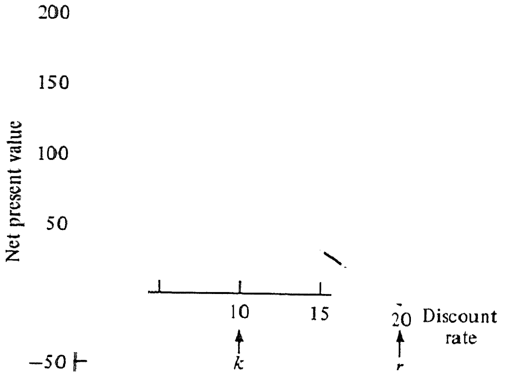


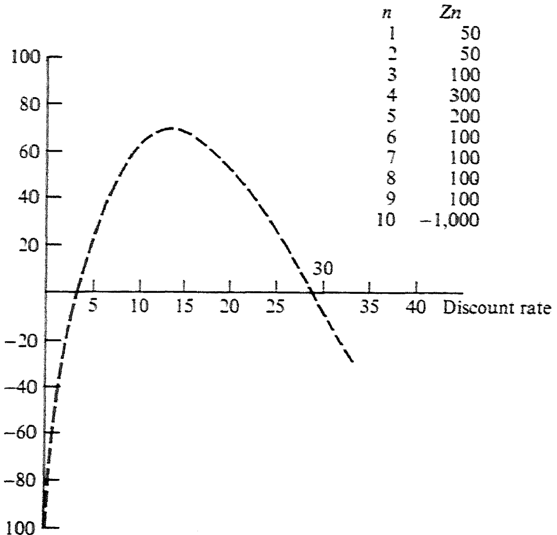
Fig. 2-1 Net present value.

investment flow to the cost of the investment, the net present value is zero at this point. Therefore the internal rate of return can be read from the net present value curve at the point where the curve crosses the discount rate axis. It can be seen that this occurs at about 19 percent. For most investments, the net present value profile will look like Fig. 2-1. The point at which the curve crosses the discount axis r will lie to the right of the cost of acquiring funds k. In addition the net present value at k will be positive. Only when both of these are true will the two methods lead to the same conclusions for a single investment.

Multiple rates of return When the net present value profile is not nearly as well behaved, the conclusions between the two methods can disagree. For example, if benefits Z_n in any future year are negative, there may be more than one internal rate of return. This situation occurs in the extractive industries such as mining, where net outflows of cash are common during the end of a project's life. The excess of expenses over revenues might stem from the closing costs associated with mines. Consider the following project, which costs 200 and has the pattern of cash flows shown in Fig. 2-2.

If the net present value profile is drawn, it becomes apparent that there are two places at which the curve crosses the discount axis. Therefore we have internal rates of return at 3 percent and 29 percent. If the cost of capital is 10 percent, a 3 percent rate of return merits rejection while a 29 percent rate of return warrants acceptance.

To overcome this dilemma we simply resort to the present value approach. We can see from the net present value profile that the net present value is positive at a cost of acquiring funds of 10 percent. Therefore if we accept the project, it will add to the profitability of the firm. It can therefore be concluded that if the pattern of cash flows is such that there are one or more reversals in



n	Zn
1	50
2	50
3	100
4	300
5	200
6	100
7	100
8	100
9	100
10	−1,000

Fig. 2-2 Net present value profile of a project with multiple rates of return.

the signs, there may be multiple rates of return. In situations like this, the present value method should be used.

Ranking of mutually exclusive investments The second situation when the present value and internal rate of return methods lead to conflicting results occurs when two mutually exclusive investments must be ranked. Mutually exclusive investments are those where the selection of one rules out the selection of other alternatives in the set. For example, suppose that you are considering the purchase of a new car. Your alternatives include a Ford, a Chevrolet, and a Plymouth. The choice of one eliminates the other two investments automatically. The three are called a *mutually exclusive set of investments.*

Consider the following two mutually exclusive investments where capital costs C are \$900 in both cases.

Investment A $C = 900$		*Investment B* $C = 900$	
n	Z_n	n	Z_n
1	100	1	900
2	200	2	100
3	500	3	100
4	800	4	100

First the investments will be ranked by the present value criterion. If the cost of acquiring funds is 10 percent, for investment A we have

$$\frac{\sum_{n=1}^{4} Z_n[1/(1+i)^n]}{C} = \frac{1,176}{900} = 1.31$$

The rank for investment B would be 1,047/900 = 1.16. Therefore by the present value method we would rank A first and then B.† However, if we look at the net present value profile in Fig. 2-3, we see that the internal rate of return method ranks the investments as B-A. The reason for this discrepancy is that the internal rate of return method assumes that cash generated from the investment is reinvested at the internal rate of return. In this case it is assumed that cash generated from investment B when reinvested will earn 23 percent. Intermediate cash flows from investment A are assumed to earn 20 percent. The present value approach, on the other hand, assumes that reinvestment will earn 10 percent or the cost of acquiring funds. Since it is nearly impossible to determine exactly what rate of return will be earned when cash is generated from investments, it is difficult to rationalize the use of the internal rate of return

---

† Since the cost of both investments is the same, the present value ranking can be observed directly from the present value profile. The profile with the highest net present value at the cost of acquiring funds receives the highest rating.

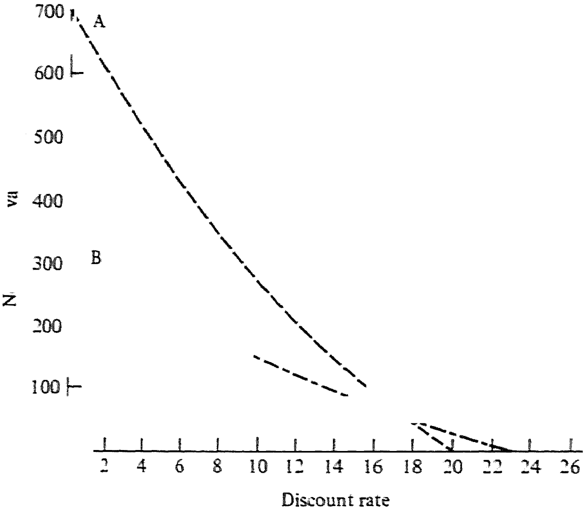


Fig. 2-3 Net present value profile for two mutually exclusive investments.

method for ranking these investments. If we use the present value method, at least we are assured that cash flows from the investment under consideration can be reinvested at the cost of acquiring funds. In fact, the chances are that we will do even better than that. Therefore, present value tends to be a conservative ranking technique but one generally considered to be theoretically superior.

A SYSTEMS APPROACH WITH LINEAR PROGRAMMING

Now that the present value approach has been settled upon as the most appropriate, a more systematic method of handling all capital expenditure proposals will be developed. Suppose that we are faced with determining which of nine alternative investments should be undertaken. The net present value of each of these investments is given as:

Investment	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*
NPV	12	15	19	11	35	10	15	9	11

Capital outflows, however, occur in both the first and the second year of the investment. In large projects this is a common situation. For example, in the first year a building is erected and in the second equipment is installed. The capital outflows for the first period and the present value of the outflows for the second period are:

Investment	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*
First period	10	32	5	6	27	5	37	32	17
Second period	2	6	5	3	27	7	12	2	2

Outflows in the first period have not been discounted, since they are incurred at the beginning of the period. The capital constraint in the first period is 50 while the present value of the capital constraint in the second period is 20. Using the benefit/cost ranking method, the present value of benefits could be divided by the present value of costs. Investments could then be selected in order of rank until capital in the first period is exhausted. This assumes, however, that sufficient capital will be available to meet the capital demands in the second period, and such may not be the case. The method that will be presented takes these multiperiod cash flows into consideration when it selects that investment mix which will maximize present value.

Our objective in building a general investment model will be to maximize the net present value of the investment mix—the value added to the firm. The choice of investments in the final mix is constrained by the capital budgets available in each of the two years. It sounds like a problem in constrained maximization. In fact, we shall first develop it as a linear programming problem.

A LINEAR PROGRAMMING MODEL

Since the objective is to maximize the net present value, the objective function can be written as

$$\text{Maximize NPV} = 12X_1 + 15X_2 + 19X_3 + 11X_4 + 35X_5 + 10X_6 + 15X_7 + 9X_8 + 11X_9$$

where X_i is called a *decision variable*. If X_i takes on the value of 1 in the final solution, the project will be undertaken. If X_i is zero, the project will not be undertaken. We will refer to the value which this decision variable takes on as the *intensity* of the project.

The capital constraints can be written for each of the two years as

$$\text{Subject to } 10X_1 + 32X_2 + 5X_3 + 6X_4 + 27X_5 + 5X_6 + 37X_7 + 32X_8 + 17X_9 \leq 50$$

$$2X_1 + 6X_2 + 5X_3 + 3X_4 + 27X_5 + 7X_6 + 12X_7 + 2X_8 + 2X_9 \leq 20$$

The first inequality states that the cash requirements for each investment multiplied by their respective intensities must not exceed the capital budget of available in that year. The second equation has a similar interpretation for the second year. Consider a *possible* solution where the following projects are undertaken:

$$\begin{array}{ll} X_1 = 1 & X_3 = 1 \\ X_4 = 1 & X_9 = 1 \end{array}$$

All other projects are rejected.

$$\begin{array}{lll} X_2 = 0 & X_5 = 0 & X_6 = 0 \\ X_7 = 0 & & X_8 = 0 \end{array}$$

Checking to see if this solution is feasible in the first budget period, we have

$$10(1) + 32(0) + 5(1) + 6(1) + 27(0) + 5(0) + 37(0) + 32(0) + 17(1) = 38$$

This is less than the budget constraint of 50; therefore the plan is feasible in the first budget period. What about the feasibility of this plan in the second period? Checking out the second constraint, we have

$$2(1) + 6(0) + 5(1) + 3(1) + 27(0) + 7(0) + 12(0) + 2(0) + 2(1) = 12$$

This is well within the capital constraint of the second period, and it can therefore be concluded that this plan is feasible. Whether or not it is optimal cannot be determined without further work. This, of course, is the job of the linear programming model; it essentially checks all these alternative sets of investments until the optimal one is found.

Specification of the model is not yet complete. Consider the situation where one of the projects is a bridge. It makes little sense for the intensity of the project to be 10, indicating that 10 identical bridges should be built. Many investment proposals are like this; at the most, you want just one of them. Incorporating this into our constraint set, we have

$$\begin{array}{l} X_1 \leq 1 \\ X_2 \leq 1 \\ X_3 \leq 1 \\ X_4 \leq 1 \\ X_5 \leq 1 \\ X_6 \leq 1 \\ X_7 \leq 1 \\ X_8 \leq 1 \\ X_9 \leq 1 \end{array}$$

Therefore each of the X's can take on a maximum value of 1. Recalling the nonnegativity conditions common to all mathematical programming models, we include

$$X_1, X_2, X_3, X_4, X_5, X_6, X_7, X_8, X_9 \geq 0$$

This completes the model. But before we discuss its solution, let us write it out, using general notation. The objective function can be written as

$$\text{Maximize NPV} = \sum_{i=1}^{n} d_i X_i$$

where NPV = net present value of investment mix
d_i = net present value of project i
X_i = level of intensity of project i

In all, there are n such projects under consideration. The capital constraints can be written as

$$\sum_{i=1}^{n} C_{i,1} X_i \leq B_1$$

$$\sum_{i=1}^{n} C_{i,2} X_i \leq B_2$$

$$\sum_{i=1}^{n} C_{i,j} X_i \leq B_j$$

where $C_{i,1}$ = capital expenditure associated with investment i in year 1
$C_{i,2}$ = present value of capital expenditure associated with investment i in year 2
$C_{i,j}$ = present value of capital expenditure associated with project i in year j
B_1 = budget constraint in year 1
B_2 = budget constraint in year 2, and
B_j = budget constraint in year j

If there are only two capital-budgeting periods that need be considered, as there were in our example, only the first two of these inequalities would be necessary for the capital constraint set. There are, then, as many capital-budget constraint equations as there are relevant budgeting periods.

Next we generalize the constraints that limit the intensity of each project to 1.

$$X_1 \leq 1$$
$$X_2 \leq 1$$

$$X_n \leq 1$$

Finally, the nonnegativity conditions

$$X_1, X_2, \ . \ . \ . \ , X_n \geq 0$$

The capital budgeting model for maximizing NPV can then be summarized as

$$
\boxed{
\begin{array}{l}
\text{Maximize NPV} = \sum_{i=1}^{n} d_i X_i \\
\text{Subject to} \\
\qquad\qquad \sum_{i=1}^{n} C_{i,1} X_i \leq B_1 \\
\qquad\qquad \sum_{i=1}^{n} C_{i,2} X_i \leq B_2 \\
\qquad\qquad \cdot \\
\qquad\qquad \cdot \\
\qquad\qquad \cdot \\
\qquad\qquad \sum_{i=1}^{n} C_{i,j} X_i \leq B_j \\
\qquad\qquad X_1 \leq 1 \\
\qquad\qquad X_2 \leq 1 \\
\qquad\qquad \cdot \\
\qquad\qquad \cdot \\
\qquad\qquad \cdot \\
\qquad\qquad X_n \leq 1 \\
\qquad\qquad X_1, X_2, \ldots, X_n \geq 0
\end{array}
}
$$

SOLUTION OF THE LINEAR PROGRAMMING MODEL

Now we can turn to the solution. First the objective function and constraints must be written in augmented form. Then the matrix is constructed and the solution process begins. The initial matrix is shown in Table 2-2, and it is

Table 2-2 Initial linear programming matrix

		12	15	19	11	35	10	15	9	11	0	0	0	0	0	0	0	0	0	0	0	
		X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8	S_9	S_{10}	S_{11}	
0	S_1	10	32	5	6	27	5	37	32	17	1	0	0	0	0	0	0	0	0	0	0	50
0	S_2	2	6	5	3	27	7	12	2	2	0	1	0	0	0	0	0	0	0	0	0	20
0	S_3	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1
0	S_4	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1
0	S_5	0	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1
0	S_6	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1
0	S_7	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1
0	S_8	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1
0	S_9	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	1
0	S_{10}	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	0	1
0	S_{11}	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	1

clear that only the very ambitious will attempt to solve it by hand. Fortunately there have been many linear programming codes written for both batch and time-share computers. The use of these codes is quite easy and a problem of the sort just developed can be solved very quickly by means of them. It should become apparent now that hand solution of any significant linear programming problem is nearly impossible. Even our example in this chapter, with nine investment alternatives, is small compared with the size of the problem one is likely to find under actual conditions.

The solution matrix for this problem is shown in Table 2-3. From this table the following conclusions can be drawn.

$X_1 = 1.0$
$X_2 = 0$
$X_3 = 1.0$
$X_4 = 1.0$
$X_5 = .0370$
$X_6 = 1.0$
$X_7 = 0$
$X_8 = .1875$
$X_9 = 1.0$
$S_1 = 0$
$S_2 = 0$
$S_3 = 0$
$S_4 = 1$
$S_5 = 0$
$S_6 = 0$
$S_7 = .9630$
$S_8 = 0$
$S_9 = 1$
$S_{10} = .8125$
$S_{11} = 0$

The net present value of this investment mix is $65.95. Projects 1, 3, 4, 6, and 9 are undertaken in their entirety, but projects 5 and 8 are fractional projects. Since it may be difficult to build .1875 of a bridge, one alternative would be to increase the budget so that all fractional projects could be undertaken in their entirety. Theoretically, however, this should not have to be done, for we should have some way to restrict the intensity of the activities to integer values of 0 and 1. In fact, we do. Recall that integer programming fills this bill very nicely. Therefore, if we solve the problem as an integer programming problem, there will be no trouble in making sense of the answer. The values of the X's will be either 0 or 1. It is, however, a property of the linear programming model that *at most* the number of fractional projects in the solution will equal the number of budgeting periods in the model. In this case we have two budgeting periods and two fractional projects. Because the fractional problem is relatively insignificant, the linear programming approach is often taken. Here is an illustration of the point made earlier that in many situations the simpler model will suffice. Nevertheless, it requires that we understand when and why these simplifications are appropriate. It is not enough to use the simplest model and hope for the best.

The variables S_1 and S_2 are the slack variables associated with the capital-budget constraint equations in periods 1 and 2. In the final solution S_1 and S_2 are equal to zero, indicating that the entire budget in both periods was used up. Slack variables 3 through 11 are associated with the constraints that limit the intensity of each project to 1. For example, $X_1 + S_3 = 1$. Therefore, when the slacks take on a value of zero, the intensity of each project must be 1. Conversely, if the slack takes on a value of 1, then the decision variable for that project must take on a value of zero. Looking at the output, slack variable 3 is zero, indicating that the project associated with slack variable 3, or project 1, is undertaken. Likewise, slack variables 5, 6, and 8 are all zero, indicating

Table 2-3 Solution for linear programming problem

	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8	S_9	S_{10}	S_{11}	
X_8	0	.81	0	0	0	0	.78	1	0	.03	− .03	− .25	0	0	− .09	0	.06	0	0	− .47	.1875
X_5	0	.22	0	0	1	0	.44	0	0	0	.04	− .07	0	− .18	− .11	0	− .26	0	0	− .07	.0370
X_1	1	0	0	0	0	0	0	0	0	0	0	1.00	0	0	0	0	0	0	0	0	1.00
S_4	0	1.0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1.00
X_3	0	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1.00
X_4	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1.00
S_7	0	− .22	0	0	0	0	− .44	0	0	0	− .04	.07	0	.18	.11	1	.26	0	0	.07	.9630
X_6	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1.00
S_9	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	1.00
S_{10}	0	− .81	0	0	0	0	− .78	0	0	− .03	.03	.25	0	0	.09	0	− .06	0	1	.47	.8125
X_9	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	1.00
	0	− .09	0	0	0	0	−7.58	0	0	− .28	−1.01	−7.16	0	−12.5	−6.26	0	−1.49	0	0	−4.19	

Note: For complete interpretation of this table, see Appendix A.

that projects 3, 4, and 6 are undertaken. Slack variables 4 and 9 are 1, since projects 2 and 7 were not undertaken. Fractional values for slack variables indicate that a project is partially undertaken. Since $X_5 + S_7 = 1$ and $X_5 = 0.37$, then $S_7 = 1.000 - .037 = .963$. In addition to slack 7, slack 10 took on a fractional value, indicating that project 8 was fractionally undertaken.

The shadow prices also provide some useful information. For example, the shadow prices associated with the capital-budget constraints indicate that if the capital budget in the first period were increased by 1, the net present value of the objective function would increase by .28. The shadow price for the capital constraint in the second period informs us that an increase in the capital budget of $1 would increase the net present value by $1.01. We might therefore conclude that a shift in funds from period 1 to period 2 might be desirable.

SENSITIVITY ANALYSIS

Perhaps the most beneficial insight into the problem can be gained by performing relatively simple sensitivity analysis. Through this analysis changes in any of the parameters can be investigated. For example, the consequence of more or less funds in any period can be determined. Or the decision maker might like to know which investment would be dropped if the budget were reduced by 20 percent. Since the present value of the projects under consideration is never known with complete certainty, it might be helpful to know the way in which changes in these net present values will alter the optimum set of investments.

It is, of course, quite helpful to be aware of the sensitivity that the model exhibits to changes in any of these factors. If the model is relatively insensitive, there is a broad range of values for these factors that will not significantly change the best set of investments. If, on the other hand, the model is very sensitive to these data, the data collection phase becomes more critical.

MUTUALLY EXCLUSIVE INVESTMENTS

Up to this point we have considered only independent investments. The extension to a set of investments including both independent and mutually exclusive alternatives is simple.

Consider that investments 1 and 3 are mutually exclusive. They might represent two types of milling machines. Only the better of the two should be selected. We can therefore add the constraint

$$X_1 + X_3 \leq 1$$

Notice that if $X_1 = 1$, then X_3 must be 0; and if $X_3 = 1$, then $X_1 = 0$. Also permissible are $X_1 = 0$ and $X_3 = 0$, but it is *not* possible for both $X_1 = 1$ and $X_3 = 1$ to be accepted.

The addition of this constraint makes the unity constraints of the investments in the mutually exclusive set redundant. We therefore can omit $X_1 \leq 1$ and $X_3 \leq 1$.

MANPOWER CONSTRAINTS

In some cases the selection of an investment set depends not only on the productivity of the investment and the capital availability, but also on the manpower availability. Some projects, although financially attractive, may make such demands on the pool of labor that the investment cannot be undertaken.

To incorporate manpower constraints into the LP (linear programming) model is relatively simple. Let L_j be the number of man-hours available in period j, and r_{ij} be the number of man-hours of input required for project i in period j. The manpower constraint for the first period can be written as

$$\sum_{i=1}^{M} r_{i1}X_i \leq L_1$$

and for the second period as

$$\sum_{i=1}^{M} r_{i2}X_i \leq L_2$$

If these constraints are added to the investment problem, then in addition to complying with the budget constraints in both periods, the optimal investment set will be assured of an adequate labor force.

THE INCORPORATION OF RISK THROUGH SIMULATION OF AN INVESTMENT PROPOSAL

The overall risk faced by the firm is related to the expected stability of the future returns on all of its investments. If the firm maintains a risky set of investments which have associated with them an uncertain future stream of income, the risk complexion of the firm can be said to be high.

Up to now we have assumed that the risk complexion of the firm would not be affected by the acceptance of any proposal. Since some proposals are inherently more or less risky than others, this may not be true. If the firm reduces or increases the riskiness of its investment mix, the risk complexion of the firm may indeed be altered and the investors may change their valuation of the firm.

Take, for example, the situation where a firm is considering an investment mix which has a higher average yield than the yield associated with its present set of investments. The risk associated with this mix, however, is much greater. It is therefore quite possible that in this shift to a riskier set of investments, the investors are likely to lower their valuation of the stock. In spite of the fact that the average *expected* yield from the new portfolio of investments is higher, this may not be enough to offset the higher risk. It becomes apparent, then, that the riskiness of an investment warrants closer scrutiny.

Exactly what the overall implications of risk are on the valuation of the firm is a bit outside the scope of this book. What is our concern, moreover, is the risk of the individual investment. Consider the example shown in Fig. 2-4 where an investment has an 80 percent chance of yielding a benefit/cost ratio of 2.5, a 10 percent chance of a 5 ratio, and a 10 percent chance of a zero ratio.

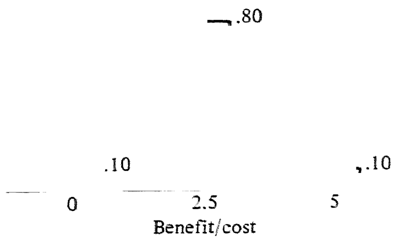


Fig. 2-4 Benefit/cost distribution.

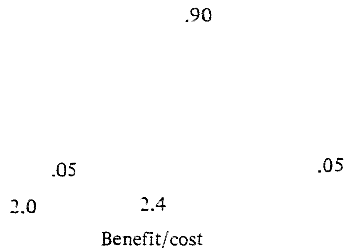


Fig. 2-5 Benefit/cost distribution.

The average benefit/cost ratio is 2.5, and it was an average figure of this sort that we were content with using earlier in the chapter. The point to be made now, however, is that investments nearly always have some degree of risk, and this risk can be measured by the variability of outcomes about the mean. Consider another investment shown in Fig. 2-5 where this time there is a 90 percent chance of yielding a benefit/cost ratio of 2.4, a 5 percent chance of a 2.8 ratio, and a 5 percent chance of a 2.0 ratio. Although the mean benefit/cost ratio of 2.4 is lower than the first investment, the variability around this mean is much less. In fact, it is certain that the benefit/cost ratio will be 2 or higher in the second investment, whereas in the first investment a ratio of 0 will occur 1 time out of 10. If you had to choose between these two investments, would the increase in benefit/cost ratio from 2.4 to 2.5 be worth this added risk? It is doubtful. The message, then, is that to base a decision on mean values alone is perhaps too superficial. The degree of risk should be considered, and it can be measured by the variability in the benefit/cost ratio about the mean. To disregard the issue of variability is tantamount to a non-swimmer attempting to wade across a stream knowing only that its average depth is 2 feet.

If the investment is large in relation to the assets of the firm, it is even easier to see why risk should be taken into consideration. Even a 10 percent chance of no return (0 benefit/cost ratio) could force the firm to reject the proposal, when a loss of this magnitude might lead to bankruptcy.

THE SOURCES OF RISK

The first step in risk analysis is to uncover the major factors which contribute to the risk of the investment. In order to keep the example at a manageable size, let us consider just three factors that can account for the variability of a particular investment. These will include the cost of the project, the life of the project, and finally its profit contribution.

Cost of the project The exact cost of a project in many cases is not known in advance. There is always the chance that the actual cost will vary from the original estimate. One can never foresee exactly what the construction, debugging, design, and developmental costs will be. Rather than being satisfied with a single estimate, it seems more realistic to specify a range of costs and the likelihood of occurrence for each value within the range. Of course, the less confidence the decision maker has in his estimate, the wider will be the range.

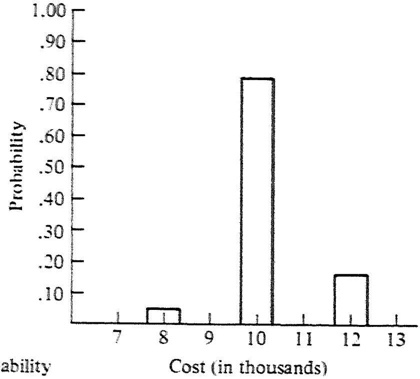
Consider the situation where a firm is about to invest in a project to manufacture a new product. The production manager feels that there is about an 80 percent chance that this project will cost $10,000; however, if things go exceedingly well and there are no training problems, it might cost as little as $8,000. He feels that the chance of this occurring is 5 percent. On the other hand, if there are either extended debugging or training periods, costs could go to $12,000. The chance of this is 15 percent. By using his intuition in developing the likelihood of possible outcomes, he has developed what we will call a *subjective* probability distribution of possible outcomes. It is shown in Fig. 2-6.

Life of the project As is always the case, the life of a project is never known with certainty. In this example, the production manager feels that the life of the project can be expressed as

Life	*Probability*
5	.10
6	.20
7	.40
8	.20
9	.10

It is further assumed that the equipment will have little salvage value at the end of its life, and this dimension can therefore be neglected.

Profitability of the project The final consideration is the revenue-minus-expense stream (Z). For simplicity assume that this stream will be the same over the



Cost	Probability
8,000	.05
10,000	.80
12,000	.15

Fig. 2-6 Investment cost.

life of the project once the level has been determined. The level can be represented by

	Probability
\$2,000	.10
3,000	.40
4,000	.40
5,000	.10

That is, there is a 10 percent chance that the stream (Z) will be \$2,000 each year, a 40 percent chance that the stream will be \$3,000 each year, etc.

PROCESS OF SIMULATION

Now that the various distributions have been specified, the process of simulating a series of benefit/cost ratios can proceed. In each simulation run five steps will be undertaken. First, we will generate a cost of the project from the cost distribution. Second, we will generate a life of the project. Third, we will generate a revenue-minus-expense stream. Then the present value of this stream will be computed at the firm's cost of capital, 10 percent. Finally the benefit/cost ratio will be calculated by dividing the present value of the benefit stream by the cost figure. The process is repeated as another sample is taken so that a range of benefit/cost ratios can be established to convey some measure of risk.

To set the stage for the simulation runs, two-digit numbers must be assigned to the probability distributions (see Table 2-4).

The first simulation run is started by generating three random numbers which are used to select a cost, life, and profit contribution from the probability distributions. The following random numbers are selected from the random number table: 74, 91, 33. These numbers correspond to a cost of \$10,000, a life of 9 years, and a profit contribution of \$3,000 per year. The present value of \$3,000 received each year for 9 years is \$17,161.20.

n	Z_n	$\frac{1}{(1+i)^n}$	$Z_n \frac{1}{(1+i)^n}$
1	\$3,000	.9091	\$ 2,727.3
2	3,000	.8264	2,479.2
3	3,000	.7513	2,253.9
4	3,000	.6830	2,049.0
5	3,000	.6209	1,862.7
6	3,000	.5645	1,693.5
7	3,000	.5132	1,539.6
8	3,000	.4665	1,399.5
9	3,000	.3855	1,156.5
			\$17,161.2

Table 2-4

Cost

C	*P(C)*	*Corresponding two-digit numbers*
$ 8,000	.05	00–04
10,000	.80	05–84
12,000	.15	85–99

Life

L	*P(L)*	*Corresponding two-digit numbers*
5	.10	0–9
6	.20	10–29
7	.40	30–69
8	.20	70–89
9	.10	90–99

Profit contribution

Z	*P(Z)*	*Corresponding two-digit numbers*
$2,000	.1	0–9
3,000	.4	10–49
4,000	.4	50–89
5,000	.1	90–99

The benefit/cost ratio is therefore 17,161/10,000 = 1.72. This benefit/cost ratio, however, is the result of a single replication of the model. Other replications would be expected to generate different values. The results of twenty-five such replications are shown in Table 2-5.

The average of these benefit/cost ratios was computed to be 1.64, and although by itself it might imply a profitable project, the use of all 25 replications will communicate the risk associated with this investment more effectively. One way to accomplish this is to construct a frequency table and then plot a frequency histogram. This is done in Fig. 2-7.

The distribution shows us not only the mean, but the likelihood that the benefit/cost ratio will go below 1 or some other value of interest. In this case there is a good chance (about 5 in 25) that the benefit/cost ratio could be below 1.30 and virtually no chance that it will be above 2.50.

There are three ways in which risk can be communicated: (1) by specifying the range of possible outcomes, (2) by presenting the actual frequency distribu-

Table 2-5 Simulation results

Run	*Benefit/cost*	*Run*	*Benefit/cost*	*Run*	*Benefit/cost*
1	2.13	11	1.83	21	1.60
2	1.02	12	2.43	22	1.60
3	1.43	13	1.46	23	0.97
4	1.72	14	1.95	24	1.95
5	1.92	15	1.31	25	2.13
6	2.30	16	1.45		
7	1.95	17	1.15		
8	1.60	18	0.72		
9	1.31	19	2.43		
10	1.07	20	1.60		

tion of outcomes, and (3) by calculating the standard deviation. The range is the easiest one to use and in this case would be .72–2.43. The frequency distribution, however, communicates more information in that one can see the likelihood of attaining each value within the range of values. Using the range, on the other hand, it is impossible to determine the likelihood of each value within the range. The standard deviation is statistically more useful, especially in more advanced analysis.

SENSITIVITY ANALYSIS

If there were some question as to the actual shape of any of these input distributions, life, cost, and revenue-minus-expense stream, other distributions could be used in their place. In this way the final benefit/cost distribution could be

Interval	Frequency
.70- .89	1
.90-1.09	3
1.10-1.29	1
1.30-1.49	5
1.50-1.69	4
1.70-1.89	2
1.90-2.09	4
2.10-2.29	2
2.30-2.49	3

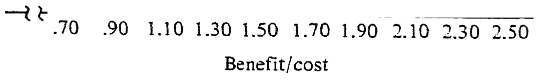


Fig. 2-7 Frequency table and benefit/cost distribution.

examined to see the way in which it is affected by changes in the input distributions.

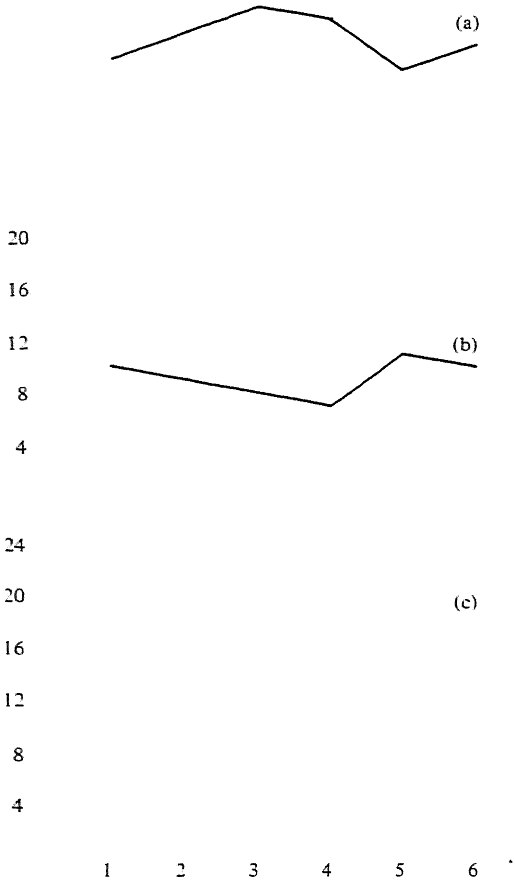
PROJECT RISK AND THE FIRM'S RISK

The relationship between the riskiness of the project, which we have just analyzed, and that of the firm is a complex one. Of primary concern is the marginal contribution to the firm's risk made by the proposed investment. The main point made in this section will be that the effect that a proposed investment will have on the firm cannot be analyzed in isolation from the firm's current set of investments. There is a significant interaction that must be taken into consideration. For example, when we combine two risky investments, it may be that the investment "portfolio" so generated will be less risky than the riskiest of the two; in some cases it will be much less. When can this happen? When fluctuations in the income stream from one investment are offset by fluctuations in the other direction for the second investment. Let us turn to an illustration of this in securities analysis. Consider two stocks with high yields. Their dividend streams historically have had wide fluctuations. Therefore, if we use this historical information to estimate their dispersion of income, both stocks can be said to be risky. We then have a mean value and a measure of the dispersion, which are all we need to evaluate the productivity and riskiness of the investments considered singly. But in *combining* investments, there is an extra dimension that should be considered. The first stock's fluctuations might coincide with the business cycle while the second stock might fluctuate in the opposite direction. That is, when dividends from the first stock have historically increased by 10 percent, dividends from the second stock have historically decreased by nearly the same percent. When both stocks are combined in a portfolio, the average yield remains the same but the risk is substantially decreased. This is precisely why investment portfolios are diversified.

Therefore, in keeping with our objective to maximize the value of the firm to its shareholders, it becomes necessary to look at prospective investment projects in the context of the present investment mix rather than by themselves. From the firm's point of view, it should see the equipment selection process as part of its job in managing an equipment portfolio. This is the same approach that a stock portfolio manager should take when either compiling a new portfolio or deciding which of many alternative stocks should be added to a present portfolio.

What is the objective in designing such a portfolio of stock (or equipment)? That for a given level of risk, you get the maximum yield (net present value), or conversely, for a given level of yield, you are subjected to the lowest possible risk.

Let us be more explicit. For simplicity assume that we have two projects, A and B. The flows from these projects are expected to be as shown and plotted in Fig. 2-8. If both projects are selected, the combined flow will be as shown in Fig. 2-8*c*. What has been gained by combining the two investments? The



Project A		Project B	
Year	Flow	Year	Flow
1	10	1	10
2	12	2	9
3	14	3	8
4	13	4	7
5	9	5	11
6	11	6	10

Fig. 2-8 Combination of risky investments.

average flow from the combination of the two investments will be equal to the sum of the averages from each one. However, the variability of the investment portfolio, as seen in Fig. 2-8*c* is much less than the variability of investment A or investment B. Therefore, by selecting these two investments, we have reduced our risk from what it would be if only one of them were selected.

Why did this happen? Because when flows from investment 1 decreased, flows from investment 2 increased. We say they are *negatively correlated.* When flows from one investment increase while those from the other also increase, we have positive correlation. When the two series move in exact proportion, we say correlation is perfect or 1; and when they move in exact proportion in opposite directions, correlation is perfect in the negative direction. If there is no relationship, then correlation is zero. The more usual case is to have correlation between these extremes. We can therefore conclude that the variability of a combination of two or more investments depends upon the movement through time of one flow in relation to the others, as well as upon the variability of each investment taken by itself.

It now appears that when the firm is considering the addition of an investment to the existing set, it should compare the mean and standard deviation of the set before and after the addition.

The precise methodology for measuring this is beyond our scope. In general, however, to estimate the mean and standard deviation of the new set, it is necessary to know the mean and standard deviation of the proposed investment, the mean and standard deviation of the existing set of investments, and the correlation between the proposed and existing investments. These are then combined to establish the mean and standard deviation of the new investment set.

Now the major problem in utilizing this approach should become apparent. In the previous sections we saw that the mean and variability of an investment could be estimated by simulation. But how are we to establish the correlation of one investment flow with another? When investments have some similarity to investments that have been undertaken in the past, these relationships can serve to estimate the correlation. Otherwise, judgment will have to be used to assess the magnitude of the relationship.

Let us see how this analysis affects the way in which we should look at portfolio decisions. Assume that the firm has a given set of investments in its portfolio and it would like to determine the consequence of adding one or more new investments. They may be risky investments by themselves, but when they are taken as part of the existing portfolio, the riskiness of the portfolio may be decreased.

The task faced by the firm is to determine the mean and standard deviation of different *combinations* of investments, where a combination represents the present set of investments plus one or more new investments. Assume that the statistical analysis has been undertaken and that we now possess the mean and standard deviation of several different portfolios. These are plotted in Fig. 2-9. Consider the combination represented by point *A*. By selecting in-

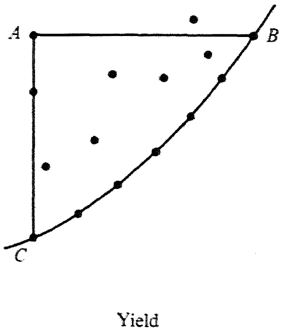


Fig. 2-9 Mean and variability of several different portfolios.

stead the combination represented by point C, we receive the same mean return but less dispersion. Therefore C is less risky at no sacrifice in yield. How about point B as contrasted with point A? With no sacrifice in dispersion, we can receive a higher yield by moving to B. In fact, all points along the line, or the "efficiency frontier," as Markowitz[2] calls it, represent the most efficient portfolios for various levels of mean yields. Where on this frontier to operate depends upon the particular preference or "utility function" of the decision maker. If he wants a higher yield, he must accept a higher risk. Once the point is chosen, the decision maker can be confident that any other combination that gives the same average yield does so at an increased risk.

In the simulation example, we looked at the riskiness of a single investment and we generated its benefit/cost ratio. However, now we have discovered that in managing a portfolio of investments, one has to look at the *net* effect of adding a project to the portfolio. The two measures of the project, its mean value and variability, by themselves cannot communicate how this investment will interact with the remainder of the investments in the portfolio.

This idea of managing the portfolio of capital investments is a relatively new one and there are many problems to be overcome in its application. In this relatively basic presentation, the objective was simply to provide some insight into the effect that each investment had on the equipment portfolio of the firm.

SUMMARY

In this chapter both theory and practice were blended to provide some insight into the way in which money is allocated among competing alternatives. Both the present value and the internal rate of return methods were presented first for the situation where capital is in unlimited supply and then for the capital-rationing case. We concluded that in general the present value method is preferred.

When the firm must consider several alternatives, each of which requires cash outflows over many periods, the use of a linear programming model can be an effective way of generating a feasible investment mix. In fact, this method can be employed whenever factors such as capital, labor, and equipment constrain the investment choice. In a large firm, with many investment proposals being generated each year, it is not hard to imagine how such a linear programming model could get exceedingly large.

It was assumed that the data required for the discounted cash flow and linear programming methods were known with certainty. This, however, is rarely true. A more general approach to the analysis of investments requires the introduction of the concept of risk. The distinguishing feature of models which incorporate this aspect is that they take into account the full range of possible outcomes for each factor rather than just its expected value. Since models of this sort generally become quite complex, simulation techniques must be employed.

---

[2] Harry M. Markowitz, *Portfolio Selection,* John Wiley & Sons, Inc., New York, 1959, p. 20.

It is also easy to see how these simulation models might become quite large. Consider the simulation example presented in this chapter. It might be reasonable to expect that more than three factors account for the variability in the benefit/cost ratio. We could include

1. Market size
2. Market growth rate
3. Share of the market
4. Investment
5. Residual value of investment
6. Operating costs
7. Fixed costs
8. Useful life of facilities[3]

As we introduce more and more variables, the necessity for using a computer becomes apparent. This is especially true if you would like the opportunity to vary some of the input distributions to determine just how sensitive the benefit/cost distribution is to changes in those factors.

Building a large-scale simulation model is therefore not very difficult. What is difficult is justifying the fact that the incremental benefit associated with a more complex model is greater than the incremental cost.

Finally, we looked at the implication that accepting risky investments had for the firm. It became apparent that certain combinations of investments could offer high benefit/cost ratios while at the same time subjecting the firm to only a modest risk. The search for these kinds of combinations or "portfolios" is familiar to security analysts but new to capital investment managers.

QUESTIONS

Q2-1. What are the three major steps in the capital-budgeting process? Explain each one.

Q2-2. Why has the payback been such a popular capital-budgeting tool?

Q2-3. What are the shortcomings of the payback method?

Q2-4. Do the benefit/cost and internal rate of return methods consistently rank investments in the same order? Explain.

Q2-5. What are mutually exclusive investments?

Q2-6. When are you likely to encounter multiple rates of return?

Q2-7. When it is impossible to allocate revenues to a general-purpose machine, what should be the basis for analyzing the investment?

Q2-8. How is the cost of capital determined?

---

[3] David B. Hertz, "Risk Analysis in Capital Investment," *Harvard Business Review*, vol. 42, no. 1, January–February, 1964, pp. 95–106.

Q2-9. A large international electrical manufacturer has a policy that investments which yield less than 15 percent should not be accepted. Where do you think they got the 15 percent, and why does the average investment, when analyzed after the fact, actually yield less than this 15 percent?

Q2-10. How might you include the eight factors specified at the end of the chapter in a large-scale investment simulation?

PROBLEMS

P2-1. The manager of manufacturing for a large shop is considering an investment with the following flows (revenue less expenses):

Year	*Flow*	*Year*	*Flow*	*Year*	*Flow*
1	\$ 0	5	\$500	9	\$400
2	25	6	500	10	300
3	100	7	500	11	200
4	200	8	500	12	100

The project will cost \$450. The payback standard for the firm is 3 years. Should he accept or reject the proposal? What, if any, shortcomings in the payback method does this problem illustrate?

P2-2. Find the present value for the following investment when the cost of capital is 15 percent.

Year	*Flow*	*Year*	*Flow*
1	\$100	5	\$200
2	200	6	100
3	300	7	50
4	300	8	−50

P2-3. Find the benefit/cost ratio for the following problem where the cost of capital is 10 percent.

Year	*Flow*	*Year*	*Flow*
1	\$10	6	\$25
2	15	7	20
3	20	8	15
4	20	9	10
5	20	10	5

A capital expenditure of $60 is incurred at the beginning of the first year.

P2-4. Find the benefit/cost ratio for the following problem where the cost of capital is 15 percent.

Year	*Flow*	*Year*	*Flow*
1	$15	5	$20
2	20	6	10
3	40	7	10
4	40	8	10

A capital expenditure of $30 must be incurred at the beginning of the first year and an expenditure of $20 at the beginning of the third.

P2-5. Projects A and B are currently being reviewed by the manager of manufacturing of Bellows, Inc. The capital cost of project A is $8,000 and that of project B is $8,000. The following is the expected stream of revenues less expenses.

Year	*Flow (A)*	*Flow (B)*
1	$1,000	$6,000
2	2,000	2,000
3	3,000	2,000
4	4,000	2,000
5	6,000	2,000

(*a*) Graph the net present value profile for each of these projects.

(*b*) What is the internal rate of return for each of them? What is their relative ranking?

(*c*) For an 8 percent cost of capital, what is the benefit/cost ratio for each project? What are their relative rankings?

(*d*) How can you account for the difference in rankings between (*b*) and (*c*)?

(*e*) Which project would you select if you only had $8,000?

P2-6. The J. M. Wall Company is a large industrial sewing shop which manufactures such garments as aluminized asbestos suits for the iron industry, heavy-duty car-top carrier covers, industrial aprons, and so on. Essentially they have several different kinds of machines through which many different jobs pass.

At present, they are considering the purchase of a new machine. The first alternative is lower in price but will incur a higher labor cost, while the second is higher in price but requires less labor. In addition, the maintenance cost for the lower-priced machine is estimated to be $1,000 for the first year, increasing by $100 per year over the life of the machine. The higher-priced machine has an estimated first-year maintenance cost of $1,500 and an increase per year of $200. The other economic data are as follows:

	Machine A	*Machine B*
Capital cost	$15,000	$25,000
Labor cost per year	14,000	7,000
Salvage value	2,000	1,000
Economic life	10 years	10 years

If their cost of capital is 15 percent, which machine should they purchase?

P2-7. The following two machines are being considered for purchase. Since they are multipurpose, it is impossible to allocate revenues to them. Which of the two should be purchased?

	Machine A	*Machine B*
Capital cost	$13,000	$8,000
Labor cost	8,000/year	9,000/year
Maintenance	300/year	400/year
Economic life	7 years	7 years
Salvage value	$2,000	$3,000
Cost of capital	10 percent	

P2-8. The Campbell Company is considering the purchase of one of two overhead cranes for their fabrication shop. For crane A the capital cost would be $25,000 while for B it would be $18,000. Both have an expected life of 10 years. Labor costs for both would be the same, but maintenance costs would not. For A, the maintenance would be $1,000 per year, while for B it would be $500 the first year, and increase by $200 each year thereafter. The salvage value for A at the end of its economic life would be $8,000 while for B it would be $3,000. If the cost of capital to the firm is 15 percent, which crane should be purchased?

P2-9. Note: The text and problems up to this point have avoided the issue of taxes. This problem gives you the opportunity to see how this dimension affects the analysis.

The Kelty Company must choose between two machines. Machine A will cost $10,000 and have a flow (revenue less expenses) of $4,000 per year for 5 years. Machine B will cost $12,000 and it will have a flow of $6,000 per year also for 5 years. Both machines have no salvage value. The company expects to depreciate the machine on a straight-line basis. It pays a 50 percent tax rate and has a 15 percent cost of capital. Which machine is preferred, and what is its after-tax benefit/cost ratio?

P2-10. A selection from among 10 investments must be made for the upcoming fiscal year. The net present value of these investments and the capital outflows for 3 years are given in the following table:

Investment	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*
NPV	25	10	14	32	16	19	37	41	14	11
Capital outflow (1)	2	17	14	12	9	16	11	12	11	6
Capital outflow (2)	7	11	14	7	11	14	11	16	2	3
Capital outflow (3)	3	6	3	15	9	2	11	2	14	3

The capital constraint for the first period is 42, for the second period it is 27, and for the third period it is 25.

(*a*) Formulate an LP model of the problem.

(*b*) How does the model change if investments 1 and 4, 7 and 9, and 8 and 10 are mutually exclusive?

(*c*) Incorporate the following manpower constraints in the model. The available manpower is 35 in year 1, 20 in year 2, and 15 in year 3.

Investment	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*
Manpower (year 1)	7	4	3	8	5	9	11	6	14	12
Manpower (year 2)	11	2	4	3	4	2	4	3	2	1
Manpower (year 3)	3	2	2	3	7	2	3	3	1	1

(*d*) How would you change the model to reflect the fact that investment 2 cannot be selected unless 3 is also selected?

(*e*) Solve the problem, including all the additional constraints, using a standard linear programming computer code. What sacrifice has to be made when solving this as an LP problem?

P2-11. John Oblanas, manager of manufacturing for Atwater Steel, would like to determine not only the profitability of a prospective investment but also its risk. He prefers not to accept it if there is more than a 5 percent chance that it will yield less than a 10 percent internal rate of return.

He subsequently estimates that the chance that the project will have a life of 10 years is 60 percent, a life of 9 years is 20 percent, a life of 8 years is 10 percent, a life of 11 years is 5 percent, and a life of 12 years is 5 percent. He also estimates that there is a 40 percent chance that the capital cost will be $10,000, a 25 percent chance that it will be $11,000, a 10 percent chance that it will be $12,000, a 10 percent chance that it will be $9,000, a 10 percent chance that it will be $8,000, and a 5 percent chance that it will be as low as $7,000. He estimates the yearly benefit stream (revenue less expenses) in the following way:

50 percent chance of $3,000, 20 percent chance of $4,000, and 30 percent chance of $2,000.

On the basis of 20 replications, should he accept the investment?

BIBLIOGRAPHY

Dean, Joel: *Capital Budgeting*, Columbia University Press, New York, 1951.

Groff, Gene K., and John F. Muth: *Operations Management: Analysis for Decisions*, Richard D. Irwin, Inc., Homewood, Ill., 1972, chap. 3.

Quirin, G. David: *The Capital Expenditure Decision*, Richard D. Irwin, Inc., Homewood, Ill., 1967.

Van Horne, James C.: *Financial Management and Policy*, Prentice-Hall, Inc., Englewood Cliffs, N.J., 1971.

Weingartner, H. Martin: *Mathematical Programming and the Analysis of Capital Budgeting Problems*, Markham Publishing Company, Chicago, 1967.

———: "Some New Views on the Payback Period and Capital Budgeting Decisions," *Management Science*, vol. 15, August 1969, pp. 594–607.

3

Replacement Analysis

INTRODUCTION

Replacement decisions are more common than new-investment decisions. In fact, they occur every day. If a machine is retained for another day, a decision has been made against replacement. Very simply, the problem is that all machines suffer from some kind of physical deterioration over time, and this in turn leads to higher maintenance costs. In addition, newer machines with lower operating costs continually emerge, and it is the combination of increasing maintenance costs and better machines that encourages replacement. If these were the only considerations in replacement, replacement would take place at frequent intervals. The missing factor is capital cost, and frequent replacement means a high average capital cost. It can then be said that replacement occurs when the operating and capital costs of the currently owned machine rise to overcome the average capital and operating costs of the replacement.

One can, then, conclude that every piece of equipment in the firm is a candidate for replacement and that operating and capital costs of the old equipment as well as operating and capital costs of the new equipment must be continually reviewed. As you will shortly see, however, there are different ways in which this analysis can be performed. In fact, the use of one model might lead to a decision quite different from that resulting from employment of another. The first set of models which we will consider come under the heading of *coterminated plant methods*. Then a model developed by Terborgh will be presented. In general, Terborgh's model is considered superior because he takes the future into account in a more explicit way than do the coterminated plant models. The coterminated plant models, however, must not be sold short. At present they are the most widely used models for making replacement decisions.

Each of these models will determine only whether a machine is a candidate for replacement. If a machine is a candidate for replacement, it must take its place with all other investment alternatives to face the final selection process. Only if its benefit/cost ratio is high enough and there is sufficient capital will it be selected. The replacement of an asset is, in most cases, no different from the consideration of a new project, and both should stand the profitability test. Computing this benefit/cost ratio for replacement decisions, however, does present some special problems. These are considered in the last part of the chapter.

There are occasions when a replacement decision does not have to take the profitability test, for example, the replacement of an X-ray unit in a hospital because it falls short of providing the kinds of services demanded by an increasingly sophisticated medical staff, or the case where a lathe is replaced because the market demands a better finish or closer tolerances. Our replacement analysis is not directed at these situations, but at replacements which are under consideration for economic reasons.

In most replacement problems the decision to continue with a project has already been made, and the piece of equipment being considered for replacement represents just one of many separate investments associated with the project, for example, a lathe in a production line or a car in a fleet of cars. In the process of evaluating the profitability of the replacement, it would be cumbersome if not altogether impossible to allocate revenues to these machines. In fact, the task of allocating expenses is not easy, but it is still more likely that the expense function can be estimated than the revenue function. Consequently the focus of interest will be on operating expenses rather than on profitability.

We will therefore assume that the profitability of the overall project of which this machine is a part was considered as a separate issue, and that the task at hand is to determine a replacement strategy for a particular machine such that we thereby maximize the productivity of the available capital. Since we will look only at the expense side of the investment, the measure of benefits which accrue from the investment will be in the form of a string of cost savings: the savings which would accrue if the new machine were in operation instead of the one presently owned. The cost element of the benefit/cost ratio will also represent a difference. This time it will be the difference between the cost of the new machine and the cost associated with keeping the present machine. For convenience, the new machine under consideration will be called the *challenger*, since it represents a challenge to the currently owned machine. It should be the best available machine. The presently owned machine, on the other hand, will be called the *defender*, since it is defending its position against the challenger.

COTERMINATED PLANT MODELS

Consider a defender with an estimated economic life of 4 years and a challenger with an economic life of 8 years. The defender, which has been in operation for some time, has a current resale value of $4,000. Since we could sell the defender and receive this $4,000, we will consider this to be the opportunity cost of continuing with the defender for 4 more years. Therefore, if we decide in favor of the challenger, the additional investment relevant for benefit/cost purposes will be the cost of the challenger, $16,000, less the resale value of the defender.

$$\begin{array}{rr} & \$16{,}000 \\ & -4{,}000 \\ \hline \text{Cost} = & \$12{,}000 \end{array}$$

The operating expenses for the defender and the challenger are given in Table 3-1. It is from these that we generate the benefit stream. The problem, however, is that when two streams are compared, selecting the appropriate number of years for comparison is made difficult if the projects do not terminate at a common period. This particular problem is referred to as *coterminated plant*. A popular approach is to choose a period of comparison in one of three ways:

Model 1. Compare the equipment over the shortest life (in this case, 4 years).
Model 2. Compare the equipment over the shortest life and include the estimated salvage or resale value of the equipment lasting longer than this period.
Model 3. Assume a common denominator for all investments (in this case, 8 years).

Table 3-1 Operating expenses

Defender		*Challenger*	
Year	*Operating expenses*	*Year*	*Operating expenses*
1	$8,000	1	$4,000
2	8,000	2	4,000
3	8,000	3	4,000
4	8,000	4	4,000
		5	4,000
		6	4,000
		7	4,000
		8	4,000

Table 3-2 Calculations for alternative 1

Year	*Savings*	*Discount factor*	*Present value*
1	$4,000	.9091	$ 3636
2	4,000	.8264	3305
3	4,000	.7513	3005
4	4,000	.6830	2732
			$12,678

FIRST MODEL

For our example, the first model suggests that we compare the difference in operating costs between the challenger and defender over a period of 4 years. At a 10 percent cost of capital, the relevant calculations are shown in Table 3-2. The benefit/cost ratio is 12,678/12,000 = 1.06. At this point it is premature to determine if this investment will actually be undertaken, for the benefit/cost ratio is now pooled with benefit/cost ratios for all types of investment alternatives. They are then ranked in order of decreasing magnitude, and from this list selection proceeds until the budget is exhausted. In our case it is highly unlikely that a benefit/cost ratio as low as 1.06 will be among those selected.

SECOND MODEL

Now let us see the effects of model 2 on the benefit/cost ratio. If we assume that the challenger would have a value of $2,000 at the end of the fourth year, then the present worth of the savings would be

$$12{,}678 + 2{,}000(.6830) = 14{,}044$$

and the benefit/cost ratio would increase to 14,044/12,000 = 1.17.

THIRD MODEL

If the third approach is used, 8 years becomes the common denominator. In the case of the defender, we assume that an expenditure of $4,000 is undertaken at the end of the fourth year, followed by a stream of $8,000 in expenses in years 5 through 8. This essentially duplicates the cost and expense pattern of the defender for 4 more years. The present value of the defender's capital cost is

$$4{,}000 + 4{,}000(.6830) = 6{,}732$$

The difference in capital costs now becomes

$$\begin{array}{rr} & \$16{,}000 \\ & -6{,}732 \\ \hline \text{Cost} = & \$\ 9{,}268 \end{array}$$

Table 3-3 Calculations for alternative 3

Year	*Savings*	*Discount factor*	*Present worth*
1	$4,000	.9091	$ 3636
2	4,000	.8264	3305
3	4,000	.7513	3005
4	4,000	.6830	2732
5	4,000	.6209	2484
6	4,000	.5645	2258
7	4,000	.5132	2052
8	4,000	.4665	1866
			$21,338

The present value of the savings stream is shown in Table 3-3, and the benefit/cost ratio becomes 21,338/9,268 = 2.30.

PROBLEM OF FUTURITY AND ECONOMIC LIFE

In the process of illustrating the use of these three alternatives, we have generated three different benefit/cost ratios: 1.06, 1.17, and 2.30. The natural question is, which is the correct ratio? From a theoretical point of view, none is correct. Consider the third alternative. In comparing the defender and challenger, we had to make some adjustments because the lives did not terminate at the same time. We duplicated the cost and expense pattern of the defender in years 5 through 8. On closer observation this may not be a reasonable way of solving the problem. After all, at the end of the first 4 years machines at least as good as and perhaps better than the challenger should be available. Why, then, assume the duplication of a defender which is already technologically inferior? It would therefore be much more realistic to introduce the cost and expense stream of the best available challenger into the analysis in the beginning of year 5. For example, this challenger might have a capital cost of $20,000, an operating cost per year of $3,000, and an economic life of 10 years. But now we have a defender's stream running into year 14. Again we have a situation with no common termination. In fact, the way this example is proceeding, the alternatives might never have a common point of termination and the differences between the cost streams might have to be compared endlessly into the future. Without making some oversimplifying assumptions, you can see that it is only by accident that the two streams will terminate in a common year. This apparently realistic problem makes the process of comparison very cumbersome. A framework for handling it will be developed shortly.

Another problem associated with our approach is that we have assumed an economic life. Why was the economic life given as 4 years for the defender and as 8 years for the challenger? What are the consequences of keeping the equipment for other than its economic life? It is these two problems, futurity and economic life, that we will consider next.

THE CONCEPT OF ECONOMIC LIFE

Our analysis in this and the previous chapter assumed that an economic life n was given. It will now be shown that the cost of a particular alternative is sensitive to the choice of economic life. That is, there is one economic life n^* which minimizes the total costs associated with an investment. The use of any other life places a handicap on that investment.

In an environment with no technological progress, one in which the machines that come out next year are no better than those of last year, economic life is simply that life when the sum of discounted capital and operating costs are at a minimum. However, in a world where machines are continually improved upon, the economic life of a machine is much more difficult to determine. At a minimum, this phenomenon which we will call obsolescence will serve to encourage earlier replacement, thereby reducing economic life.

A DECISION TREE APPROACH TO ECONOMIC LIFE

To see the magnitude of the economic life problem, a decision tree is drawn in Fig. 3-1. Each branch represents a decision period. In period 1 the decision can be to retain the defender D or to replace with the best available challenger C_1 in period 1.

In period 2, either the machine which is currently owned in that period can be retained or a new challenger C_2 can be purchased. Therefore the following sequences are possible by the end of the second period.

C_1, C_1	D, C_2
C_1, C_2	D, D

The tree can be carried out indefinitely, and each path through the tree represents one replacement strategy through time. Of course, with such a large tree, there are many, many paths through it.

For each one of these paths, present values for all relevant costs can be computed. For example, in the branch C_1, C_1, C_1, C_4, C_4, . . . , the best available challenger is purchased in the first period and held for three periods. The relevant costs here would be the present value of the capital cost plus operating costs over the three periods. To this is added the time-adjusted capital cost of the best available challenger which is purchased in period 4, plus the related operating costs for its life, followed by the time-adjusted capital and operating costs of the next replacement, and so on. The sum of these present values, then, represents the cost of that particular replacement strategy.

Notice also how technological obsolescence enters into the analysis. Since the parade of technological progress is endless, better machines are always being developed. It is not, however, until a replacement has taken place that the firm can realize these benefits. When, in fact, this does occur, the usual phenomenon is for operating costs of the new machine to be lower than those of the old machine when it was new.

If it would be possible to determine the present value of each path through

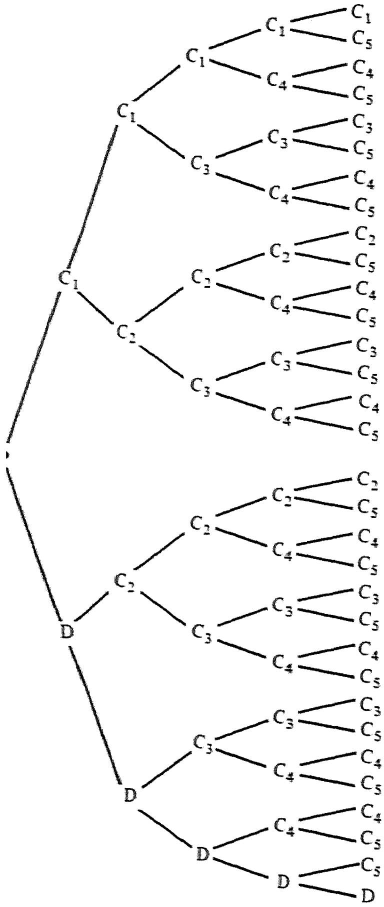


Fig. 3-1 Decision tree of replacement alternatives.

the tree, the path with the lowest value would be chosen as the best strategy. This path would have associated with it a series of machines with certain "lives" that would represent their economic lives. If any machine in the series were kept for other than its economic life, this would represent a suboptimal strategy with higher costs. The most practical piece of information is whether this chain is headed by the challenger or defender, because this tells us what immediate action should be taken: replace or keep the defender. Unfortunately we have just discovered that, to know whether this optimal chain starts with the present challenger or defender, we must evaluate every path through the tree. As it stands now, this can hardly be called an operational procedure.

THE ROLE OF THE FUTURE IN MAKING CURRENT DECISIONS

From this framework it should now be clear that the future does affect the replacement decision. As a very intuitive illustration of this, consider the situation where a new machine substantially better than anything available now will be marketed in 2 years. Not only is it expected to produce a product of higher quality, but its operating costs will be much lower than those of machines which are currently available. It might be that the best strategy is not to replace right now but to wait 2 years for this improved machine. If, however, only the present challenger was compared with the defender, it might be that replacement would have been strongly signaled. In summary, economic life involves the future, and there is no determinable relationship between the defender and the present challenger that alone and by itself can signal replacement. Somehow, then, the future must be considered.

Given the decision tree extending endlessly into the future, a solution for economic life seems hopeless. This can be attributed to the impossibility of looking ahead and determining the capital and operating costs of machines yet to be born. Secondly, even if we could estimate these costs, an infinite number of branches are quite a few to evaluate. It is therefore apparent that any solution to this problem requires some simplifying assumptions about the future. The most widely accepted have been Terborgh's assumptions, to which we will turn shortly.

FUTURE DECISIONS

Once computed does the economic life of the present challenger, as well as its successors, commit us to a procurement plan extending into the future? Absolutely not. The future will usually unfold differently from what is expected. Therefore a new study with new forecasts must be undertaken periodically. Take the case where a computer has been in operation for 5 years. When the original economic analysis was performed 5 years ago, the economic life of the computer was expected to be 8 years. A recent analysis, however, shows that the machine should be replaced now. What is the point being made? Equipment decisions are ex ante decisions. They are made before the future unfolds. Once a decision is made, this does not preclude economic analysis over the expected life of the machine. New developments which were impossible to predict accurately are always occurring. As a result, a periodic review is necessary. In this case a periodic review was made every year, and in the fifth year replacement was signaled.

TERBORGH'S MODEL

Terborgh[1] laid the foundation for an analytical analysis of the replacement problem. Before we plunge into his framework, the concepts of obsolescence and deterioration must be developed.

DETERIORATION

Deterioration is the decline in operating performance as compared with the performance that could be obtained from a new machine identical to the present one. Another way to say this is that deterioration is the increase in maintenance costs over what these costs would be with an identical new machine. The behavior of this cost is shown in Fig. 3-2. The solid line represents the operating cost of the machine and is shown rising as a result of increasing maintenance costs. When replacement takes place, the line drops down to reflect the improvement in operating costs associated with a new machine.

---

[1] George Terborgh, *Dynamic Equipment Policy*, Machinery and Allied Products Institute, Washington, D.C., 1949.

OBSOLESCENCE

The second term to be defined is *obsolescence.* Obsolescence is the decrease in operating cost associated with the best new machine over the cost associated with an identical replica of the present machine brand-new. This decrease in operating cost comes about as the result of technological progress, where better machines are developed each year. Graphically, the obsolescence factor can be seen in Fig. 3-2. It essentially represents an opportunity cost which is incurred if the best machine is not used. Combining the effects of obsolescence and deterioration, we have an inferiority gradient. The gradient will then show the increase per year in operating cost and opportunity cost due to deterioration and obsolescence.

FIRST ASSUMPTION

Next it is necessary to make two assumptions which will render our cumbersome problem into one which is manageable. The first assumption states that the "present challenger will accumulate operating inferiority at a constant rate over its life."[2] Consider a defender whose next year's operating cost is estimated to be $7,000. The best available challenger will have a next year's operating cost of $5,000, and this $2,000 difference can be attributed to the accumulation of deterioration and obsolescence over the 10 years during which the defender

---

[2] Ibid., p. 65.

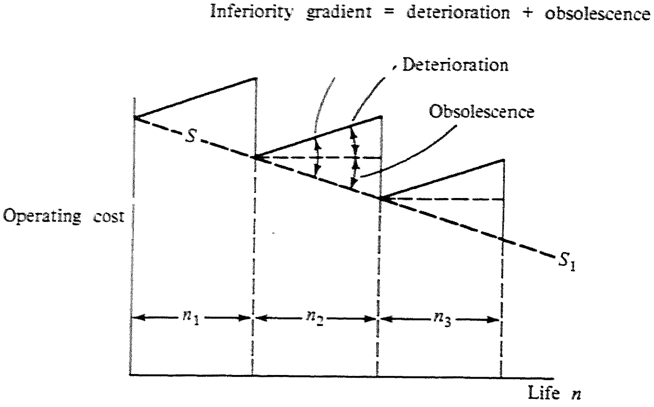


Fig. 3-2 Operating cost.

has been in service. The average increase in operating inferiority over the last 10 years can then be set at $200 per year. What is suggested by this assumption is that this average figure be used to estimate the way in which operating inferiority will accumulate over the life of the new challenger. Is it reasonable to expect that future accumulation of maintenance and obsolescence will resemble the way in which it occurred over the past? In addition, is it reasonable to assume that it will increase at a *uniform* rate of $200 per year? Perhaps it would be more realistic to assume that obsolescence occurs, not uniformly over time, but in steps. It is very possible to go for a year or more before a substantially new and better machine becomes available. When it does become available, operating costs drop instantly. Our experience tells us that this is the way things happen; however, our ability to predict this type of operating cost function is usually very limited. The reason why we return to our uniform annual increase based on past behavior is that, in the absence of a crystal ball, it seems a reasonable thing to do.

ECONOMIC LIFE OF CHALLENGER

It is now possible to determine an economic life for a single challenger, while temporarily avoiding the problems of futurity. The two sets of costs that are affected by the replacement decision include the cost of operating inferiority and the capital costs. Moreover, they move in opposite directions. As the life of the equipment increases, the average operating inferiority increases, whereas the average capital costs decrease. Adding both of these costs together, a life will be found where the sum is at a minimum. This will be the economic life. If we assume that the best available challenger costs $10,000 and that the inferiority gradient, as just determined, is $200 per year, we can then proceed to determine this life.

In Table 3-4 we find the calculations necessary to determine economic life. In column 1 is found the year of service and in column 2 the operating inferiority that would be associated with a specified year of service. For example, in the fifth year of service, the challenger suffers $800 worth of inferiority. Since these operating costs occur in the future, it is necessary to discount them to the present. Therefore the discount factors at 10 percent are multiplied by the operating inferiority, and the product is found in column 4.

The entry in column 4 represents the present worth of operating inferiority for a particular year, but what is of importance is the operating inferiority that will be incurred up to and including that particular year. For example, if the challenger is kept for 5 years, it would accumulate the present worth equivalent of $1,372 of operating inferiority.

$ 165
300
410
497
$1,372

Table 3-4

(1) Year of service	(2) Operating inferiority for year indicated	(3) Present worth factor PWF	(4) (2) × (3) Present worth of operating inferiority for year	(5) Present worth of operating inferiority for period ending with year	(6) Capital recovery factor	(7) (5) × (6) Time-adj. avg. oper. infer. for period ending with year	(8) 10,000 × (6) Time-adj. avg. cap. cost for period ending with year	(9) (7) + (8) Unif. annual equiv. for period ending with year
1	$ 0	.909		$ 0	1.100	$ 0	$11,000	$11,000
2	200	.826	$165	165	.576	95	5,760	5,855
3	400	.751	300	465	.402	187	4,020	4,207
4	600	.683	410	875	.315	275	3,150	3,425
5	800	.621	497	1,372	.264	363	2,640	3,003
6	1,000	.564	564	1,936	.230	446	2,300	2,746
7	1,200	.513	616	2,552	.205	524	2,050	2,574
8	1,400	.466	652	3,204	.187	600	1,870	2,470
9	1,600	.424	678	3,882	.174	675	1,740	2,415
10	1,800	.385	693	4,575	.163	745	1,630	2,375
11	2,000	.350	700	5,275	.154	812	1,540	2,352
12	2,200	.319	702	5,977	.147	877	1,470	2,349*
13	2,400	.290	696	6,673	.141	940	1,410	2,350
14	2,600	.263	684	7,357	.136	1,000	1,360	2,360
15	2,800	.239	669	8,026	.131	1,050	1,310	2,360
16	3,000	.218	654	8,680	.128	1,110	1,280	2,390
17	3,200	.198	634	9,314	.125	1,160	1,250	2,414
18	3,400	.180	612	9,926	.122	1,210	1,220	2,430
19	3,600	.163	587	10,513	.120	1,260	1,200	2,462
20	3,800	.149	566	11,079	.117	1,290	1,170	2,466

Column 5 represents this accumulation of operating inferiority for the various years under consideration.

If the machine is kept 8 years, we can then say that it will have accumulated the present value equivalent of $3,204 of operating inferiority. There is, however, another way to say the same thing. This other way is to express the present worth equivalent of $3,204 as an average amount which is incurred every year over the 8-year time horizon. To do this, the present worth equivalent is multiplied by a capital recovery factor which is taken from Table B-4 in Appendix B. In this example at 10 percent interest and an 8-year life, the capital recovery factor is .187. Capital recovery factors (CRF) for other lives can be found in the table and are in turn entered in column 6 of Table 3-4. Now we can determine the uniform annual equivalent (UAE) of $3,204 by multiplying $3,204 × .187 = $600. In general notation, what we have done is the following:

$$\begin{aligned} \text{UAE} &= \text{CRF}^{8}_{.10} \times \$3{,}204 \\ &= .187 \times 3{,}204 \\ \text{UAE} &= 600 \end{aligned}$$

The following interpretation can be placed on the UAE of 600: if $600 were incurred each year for 8 years, the present value of this stream at 10 percent would be $3,204. Since the $600 is a uniform amount, it is also referred to as an *annuity*. In column 7 is found the uniform annual equivalents of operating inferiority for the years under consideration.

Next we examine the behavior of the other cost that is affected by economic life: capital cost. Here we simply multiply the capital cost of $10,000 by the capital recovery factor found in column 6. Therefore, if we keep the challenger 8 years, we subject ourselves to a uniform equivalent cost of $1,870 per year. Again, we can interpret this as the uniform cost which, if incurred every year over an 8-year period, would have a present value of $10,000.

Now that the uniform annual equivalents for operating inferiority and capital costs have been calculated, they must be added together. The sum is found in column 9. We can see from this column that if the challenger were kept for 8 years, the uniform annual equivalent would be $2,470. That is, if it were kept for 8 years, the average yearly time-adjusted capital cost and operating inferiority would be $2,470. Our objective is to find the life which has associated with it the lowest uniform annual equivalent. In this case the lowest is $2,349 per year with a life of 12 years. We can therefore conclude that if the challenger is kept 12 years, it will incur a lower average capital and inferiority cost per year than for any other life. The challenger therefore has an economic life of 12 years. In addition, we call the uniform amount associated with the economic life the *adverse minimum* of the challenger. In this case the adverse minimum is $2,349. In Fig. 3-3 the relevant costs are

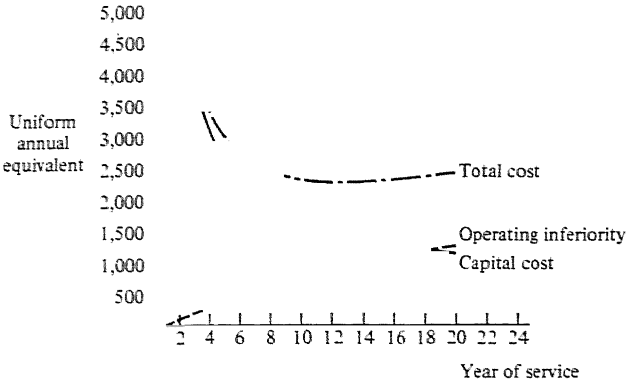


Fig. 3-3 Uniform annual equivalent.

plotted against life, n. Economic life n^* is the point at which the total cost curve reaches a minimum.

//// **General form for uniform annual equivalent** The uniform annual equivalent in any year T can be expressed as:

$$\text{UAE}(T) = C(\text{CRF}_i^T) + \left[\sum_{t=1}^{T} g(t-1)(\text{PWF}_i^t)\right] \text{CRF}_i^T$$

where C = equipment cost
CRF_i^T = capital recovery factor at i percent for T years
g = inferiority gradient
PWF_i^T = present worth factor at i percent and T years

The value of T which produces the lowest uniform annual equivalent is the economic life. ////

SECOND ASSUMPTION

We are now in a position to continue our consideration of futurity which was abandoned temporarily so that the adverse minimum could be defined. If we make a second assumption that *all future challengers have the same adverse minima*,[3] we are absolved from the problem of looking at an infinite number of branches which are infinitely long.

If in fact all future challengers have the same adverse minima, then the best succession of machines which is headed by the present challenger is a chain of machines with like adverse minima. Therefore, in this example we assume that all future machines have an adverse minimum of $2,349. This implies, of course, that each machine is replaced when it reaches the end of its economic life. As a consequence, the ~~optimal path headed~~ by the challenger has associated with it an infinite stream of $2,349.

---

[3] Terborgh, op. cit., p. 64.

For the challenger, the economic life is attained in 12 years. However, our second assumption does not require that future challengers have the same economic life. All it says is that the adverse minima will be the same, and from this we assume that the equipment will be kept for the necessary time to reach this economic life. Therefore, we cannot trace our chain past the first 12 years, since we do not know the economic life of future machines. As we will see shortly, the knowledge that the optimal chain is a series of like uniform annual equivalents is enough to absolve us from the problem of looking indefinitely into the future.

Now to the succession *headed* by the *defender*. If we look at the best succession of machines *after* the replacement of the defender, we will see a sequence of machines with like adverse minima. This, of course, is true because we assumed that all future challengers will have the same adverse minima. This means that any difference between the average adverse minimum of the chain headed by the challenger and the chain headed by the defender must occur in the period *before* the defender's replacement. Therefore it becomes necessary to compare only the adverse minimum of the defender with that of the challenger.

Let us summarize how we resolved the problem of casting indefinite comparisons into the future. We said that the challenger heads a chain of machines whose adverse minima are the same. In this case $2,349 per year is assumed to extend indefinitely into the future. *Once* the *defender* is *replaced*, we said that the uniform annual equivalent associated with the optimal branch extended into the future can also be represented by $2,349 per year, since *all* challengers will have the same adverse minima.

Therefore, the only difference between the chain headed by the challenger and that headed by the defender must occur in the period before the defender is replaced. Since the adverse minimum of the challenger has already been computed, it now remains to compute this figure for the defender.

ECONOMIC LIFE OF DEFENDER

In computing the adverse minimum for the defender, it will be assumed that the defender has a salvage value of $2,000 in the current period. The schedule of salvage values for future periods can be found in column 2 of Table 3-5. Recall that in the case of the challenger, salvage values were assumed to be zero. Therefore in calculating the adverse minimum of the defender, the only change that will take place in our methodology will be in the way we handle salvage values. This modification could also be incorporated into the calculation of the challenger's adverse minimum in the event that salvage values apply.

In the presence of salvage values, the capital cost of holding the investment for another year can be calculated as the loss in salvage value during that year plus the interest on the *opening* salvage value. This represents the oppor-

Table 3-5 Adverse minimum of defender

(1) Year of further service	(2) Salvage value at end of year	(3) Loss in salvage value for year	(4) Interest on opening salvage value for year	(5) Capital cost for year (3) + (4)	(6) Present worth factor (PWF)	(7) Present worth of cap. cost for year (5) × (6)	(8) Present worth of cap. cost for per. end. with year	(9) Capital recovery factor	(10) Time-adjusted average cap. cost for per. end. with year (8) × (9)	(11) Operating inferiority for year	(12) Present worth of oper. infer. for year (11) × (6)	(13) Present worth of oper. infer. for period end. with year	(14) Time-adj. aver. oper. infer. for per. end. with year (13) × (9)	(15) Unif. annu. equiv. for per. end. with year (10) + (14)
0	2,000													
1	1,000	1,000	200	1,200	.909	1,091	1,091	1.100	1,200	2,000	1,818	1,818	2,000	3,200
2	400	600	100	700	.826	578	1,669	.576	961	2,200	1,817	3,635	2,094	3,055
3	0	400	40	440	.751	330	2,000	.402	804	2,400	1,802	5,437	2,186	2,990
4	0	0	0	0	.683	0	2,000	.315	630	2,600	1,776	7,213	2,272	2,902*
5	0	0	0	0	.621	0	2,000	.264	528	2,800	1,739	8,952	2,363	2,891
6	0	0	0	0	.564	0	2,000	.230	460	3,000	1,692	10,644	2,448	2,908

tunity cost of holding the investment for another year. In the first year the loss in salvage value is $1,000 and the interest on the opening salvage value is $2,000(.10) = $200. See Table 3-5. The capital cost for the year is the sum of these costs, or $1,000 + $200 = $1,200. The operating inferiority for the first year is $2,000. Recall that it was this level of operating inferiority which led to the calculation of the inferiority gradient. We had established that since it took 10 years to accumulate an inferiority gap of $2,000, the average annual accumulation of inferiority was $200 per year.

In the second year the loss in salvage value plus interest on opening salvage value amounts to

$$600 + 1{,}000(.10) = 700$$

If the defender is kept for 2 years, it will accumulate a present value capital cost of $1,091 in the first year and $578 in the second (see column 7). At the end of the second year it will therefore have accumulated a total present value of $1,669 in capital costs (column 8). This is equivalent to a uniform average annual capital cost of $961 per year for each of the 2 years (column 10). The total present value of operating inferiority that is accumulated by the end of the second year is 1,818 + 1,817 (see column 12). Its sum is $3,635 and is entered in column 13. The uniform average annual inferiority cost is then computed in column 14. It is $2,094. Finally the uniform average capital and inferiority costs are added in column 15. We can conclude that if the defender is kept for 2 years, its uniform average total cost will be $3,055 per year.

Calculations are shown for 1 to 6 years, and it becomes apparent that the adverse minimum of the defender occurs in the fifth year. The replacement decision, as we determined, hinges on the comparison of the adverse minimum of the challenger with that of the defender. Since the challenger's minimum is $2,349 and the defender's is $2,891, replacement appears profitable. This conclusion, however, does not take into consideration the productivity of the capital if it is used to replace the defender. Only if capital is unlimited is it reasonable to replace when the adverse minimum of the defender is lower than that of the challenger. What we need is a benefit/cost ratio so that replacement alternatives can take their place beside all other kinds of investment alternatives in the final selection process. It is to this problem that we turn shortly. Before we do this, however, a shortcut method for finding the adverse minimum of both the challenger and the defender will be presented.

TERBORGH'S SHORTCUT

Terborgh also developed a shortcut method which is very useful in some situations. First we look at the shortcut for determining the challenger's economic life and adverse minimum, and then at the defender's.

SHORTCUT: CHALLENGER

Approximations to the adverse minimum of the challenger can be made by using a simple formula. The results turn out to be quite good. The uniform annual equivalent for the period ending with year n can be approximated as[4]

$$\text{UAE} = \frac{g(n-1)}{2} + \frac{C-S}{n} + \frac{i(C+S)}{2} \tag{3-1}$$

where UAE = uniform annual equivalent for period ending with year n
g = inferiority gradient
C = capital cost
S = salvage value in year n
i = interest rate

The first term $g(n-1)/2$ represents an approximation to the time-adjusted annual average of operating inferiority in year n. In our example it would estimate this for year 5 as

$$\frac{200(5-1)}{2} = \$400$$

From Table 3-4 we can read that the actual amount is \$363.

The next two terms represent an approximation to the time-adjusted annual average of capital costs in the year n. In our example the fifth-year estimate would be

$$\frac{C-S}{n} + \frac{i(C+S)}{2}$$

$$\frac{10{,}000}{5} + \frac{.10(10{,}000)}{2} = 2{,}500$$

The actual figure from Table 3-4 is \$2,640. Combining these two estimates, we have UAE = \$400 + \$2,500 = \$2,900. This is very close to the actual value of \$3,003.

The question is, For what value of n is Eq. 3-1 at a minimum? Once this question is answered, we know the economic life and the adverse minimum. If we simplify Eq. 3-1 so that salvage value is zero, we have a function in one variable n.

$$\text{UAE} = \frac{g(n-1)}{2} + \frac{C}{n} + \frac{iC}{2}$$

---

[4] For a complete development of the shortcut, see Floyd Gillis, *Managerial Economics*, Addison-Wesley Publishing Company, Reading, Mass., 1969.

Using calculus the life n^* which minimizes this expression can be found to be[5]

$$\sqrt{2C}$$

Returning to our example, the economic life of the challenger is

$$\frac{\sqrt{2C}}{g}$$

$$n^* = \sqrt{\frac{2(10{,}000)}{200}} = 10 \text{ years}$$

This compares with 12 years we calculated using the exact approach. The adverse minimum can now be calculated

$$\begin{aligned}
\text{UAE}^* = \text{AM} &= \frac{g(n^* - 1)}{2} + \frac{C}{n^*} + \frac{iC}{2} \\
&= \frac{200(9)}{2} + \frac{10{,}000}{10} + \frac{.10(10{,}000)}{2} \\
&= 900 + 1{,}000 + 500 \\
\text{AM} &= 2{,}400
\end{aligned}$$

This estimate compares with \$2,347 found with the exact approach.

SHORTCUT: DEFENDER

It usually happens that the defender has been in service for many years before the replacement decision is considered. Therefore it has been argued[6] that a reasonable assumption is that the adverse minimum can be approximated by the next-year sum of capital cost and operating inferiority. The implication

---

[5] The minimum of such a function can be found by taking its first derivative and then determining where this derivative is equal to zero. Taking the derivative, we have

$$\frac{d(\text{UAE})}{dn} = \frac{g}{2} - \frac{C}{n^2}$$

Setting equal to zero, to find the value of n for which the function is at its minimum,

$$\frac{g}{2} - \frac{C}{n^2} = 0$$

$$\frac{g}{2} = \frac{C}{n^2}$$

$$n^2 = \frac{2C}{g}$$

$$\frac{\sqrt{2C}}{g}$$

[6] Terborgh, op. cit., p. 120.

is that the time-adjusted cost is already rising from year to year, therefore the value for next year will be lower than for any other. The comparison under this short-cut method is:

AM of challenger	*AM of defender*
2,400	$2{,}000 + 1{,}200 = 3{,}200$

The decision, then, is the same under both methods: replace.

In our shortcut formula for economic life we assumed that salvage value for the challenger was zero. What happens if salvage value is greater than zero? The adverse minimum for the challenger will be even lower. Therefore if replacement is signaled without the consideration of salvage values, the signal would have been even stronger had salvage values been included. We can therefore conclude that if the shortcut analysis is undertaken with the assumption that salvage values are zero, and if replacement is signaled, then it is unnecessary to go further: replacement should take place. If, on the other hand, replacement in the "no salvage" case is not signaled, then different values of life must be tried in Eq. 3-1 until trial and error shows that the minimum value has been reached. Terborgh states, however, ". . . that if salvage values are not effective during the first 5 years of the challenger's service life they are most unlikely, even if effective later on, to produce an adverse minimum significantly below the no-salvage minimum."[7]

//// BENEFIT/COST RATIO FOR REPLACEMENT ALTERNATIVES

The economic-life approach can be extended to cover situations of capital rationing. In keeping with the approach developed in the first part of this chapter, we will proceed to generate a benefit/cost ratio for the problem just presented. Remember that the benefits represent the savings in costs associated with investing in the challenger. The costs represent those increases in capital expenditures associated with the challenger over the costs of the defender. Let us assume that the two alternatives we face are: (1) replace the defender immediately, (2) replace the defender at the end of the fifth year, the end of its economic life. In order to make the mathematics of the analysis manageable, some additional assumptions will have to be made. In addition to the adverse minimum of all future challengers being the same, the *life* of the challengers and their *capital costs* will also be assumed to be identical. Therefore, if we choose the first alternative and replace the defender immediately, we will hold the challenger 12 years, replace it, keep the next new one 12 years, replace it, and so on.

In the second alternative, we assume that the defender is kept for 5 years, at the end of which it is replaced by the best available challenger. This challenger, which is 5 years better, in terms of obsolescence, is kept for 12 years, replaced, the replacement kept for 12 years, replaced, and so on. To write out the stream of operating expenses, it is necessary to know the initial level of operating expenses for the challenger. This information was not necessary in the economic-life approach, since we were interested only in the *change* in inferiority costs over time g. Therefore, let us assume that the operating cost associated with the challenger in the first year of operation is $5,000. In addition, it is necessary to know the breakdown of deterioration and obsolescence in the inferiority gradient. Assume it to be $100 for deterioration and $100 for obsolescence. We can now write out the stream of expenses for each alternative.

---

[7] Ibid., p. 101.

EXPENSE STREAM: ALTERNATIVE 1

First, consider the alternative where the defender is replaced immediately. Operating costs will be $5,000 in the first year. In the second year operating costs increase by $100 to $5,100. This increase can be attributed to the increase in maintenance costs, which we have labeled "deterioration." In the next year, costs increase by another $100 to $5,200. Operating costs increase in this manner until the beginning of the thirteenth year, at which time replacement occurs. It is important to realize that the $100 increase in operating costs per year can be credited to deterioration. Obsolescence has been quietly accumulating at the rate of $100 per year for 12 years. However, no impact from this phenomenon is felt on operating costs until a new machine is placed in operation. That is, every year machines are getting better, but it is only when replacement occurs that this benefit can be realized. In our example, operating costs would drop to $3,800 when replacement occurs at the beginning of year 13. This is $1,200 lower than the operating cost of the old challenger when it was *new*. A graphic illustration of this phenomenon is presented in Fig. 3-4. The cost function increases every year until the new challenger is purchased, at which time it drops down to an operating cost associated with the best possible challenger. In year 14, maintenance costs resume their increase of $100 per year from the new base of $3,800.

Year	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*
Operating cost	5,000	5,100	5,200	5,300	5,400	5,500	5,600	5,700
Year	*9*	*10*	*11*	*12*	*13*	*14*	*15*	*16*
Operating cost	5,800	5,900	6,000	6,100	3,800	3,900	4,000	4,100

EXPENSE STREAM: ALTERNATIVE 2

Now consider the pattern of costs for the second alternative, where the defender is kept for its economic life of 5 years and then replaced with the best available

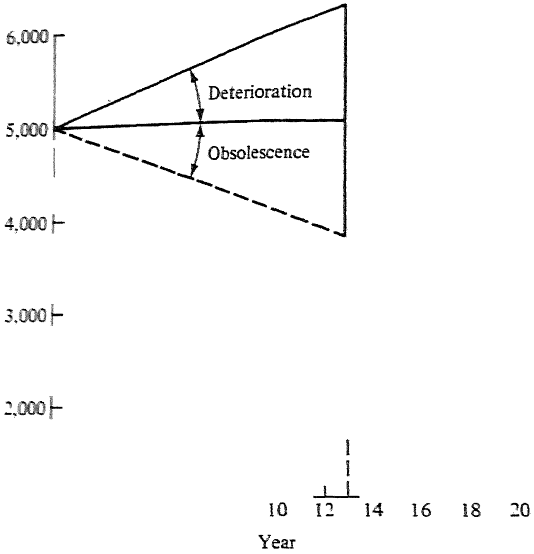


Fig. 3-4 Operating cost function.

challenger in the beginning of the sixth year. In the first year, operating costs are $7,000, since we have already established that the operating inferiority of the defender in this year is $2,000. In fact, we used this operating inferiority that was accumulated over 10 years to arrive at an inferiority gradient of $200 per period.

Year	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*
Operating cost	7,000	7,100	7,200	7,300	7,400	4,500	4,600	4,700
Year	*9*	*10*	*11*	*12*	*13*	*14*	*15*	*16*
Operating cost	4,800	4,900	5,000	5,100	5,200	5,300	5,400	5,500

In the beginning of the sixth period, we replace with the best available challenger. The operating cost associated with this challenger is $4,500. We get this by reducing the operating cost of the best available challenger in period 1 by $500, the obsolescence which has accrued over 5 years. Again the effects of deterioration cause operating costs to climb by $100 per period until the challenger is ready for replacement. Since we made the assumption that the economic lives of all future machines must be identical, we keep this challenger for 12 years. Therefore in year 18 replacement with the new best available challenger is undertaken, and since we have assumed that obsolescence costs are predictable, the new operating cost in year 18 will be \$4,500 — \$1,200 = \$3,300, where \$4,500 represents the operating cost of the old challenger and \$1,200 is the improvement which has been generated by technological progress or what we have termed "obsolescence." Again after this replacement we see costs rising by \$100 per year for the next 12 years, followed by replacement, and so on.

BENEFIT STREAM

In determining the benefits which are associated with this investment, it is the stream of savings that accrue from investing in the challenger that is relevant. Remember that replacement buys a stream of savings and that this represents the benefit. The additional outlay required to purchase the challenger represents the cost, and from these two we can generate a benefit/cost ratio. Our next task, then, is to formulate this stream of benefits and then the stream of costs. First the benefit stream.

The differences in the benefit stream are presented in Table 3-6. It is apparent that a pattern is taking place: that is, a savings in the first 5 years because a new challenger is in operation, followed by a string of losses until replacement takes place in year 13. The losses can be attributed to the fact that in the sixth year the defender is replaced with the best available challenger, one which has operating expenses \$500 lower than the operating expenses of the one that was available in period 1.

This stream, which includes five periods of savings followed by seven periods of losses, repeats from years 13 through 24, years 25 through 36, and so on. The only irregularity in this pattern is found in the first 12-year sequence. For here we have a string of \$2,000 gains followed by a string of \$1,000 losses. In the next sequence we have a string of \$1,400 gains followed by a string of \$1,000 losses, and so on into the future, with a series of \$1,400 gains followed by \$1,000 losses. Therefore, in order to measure the benefit associated with investing in the challenger now versus waiting 5 years, consideration must be given to this stream into the future. Fortunately, the discounting of the flows makes this task manageable.

Taking into account the time value of money, we can write the present value of this stream. The approach taken will be to look at each 12-year sequence. Starting with the first, the present value can be written as

$$2{,}000[\mathrm{PW}(U)^{5}_{.10}] - 1{,}000[\mathrm{PW}(U)^{7}_{.10}][\mathrm{PW}(S)^{5}_{.10}]$$

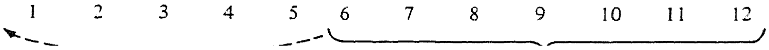
The first term, $2{,}000[\mathrm{PW}(U)^{5}_{.10}]$, determines the present worth of a uniform stream of \$2,000 gains for 5 years at 10 percent. Rather than take the present value of each \$2,000 sum over the 5 years individually, it is much more convenient to use an annuity table. Annuity tables can be used *only* when the flows are *identical* from one year to the next. In this case, the flow of \$2,000 is identical for the first 5 years. The table tells us the present value of a dollar received annually for n years. This can be written as $\mathrm{PW}(U)^{n}_{i}$, where i is the relevant interest rate.

Table 3-6 Stream of benefits

Year	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*
Alternative 1	5,000	5,100	5,200	5,300	5,400	5,500	5,600	5,700	5,800	5,900	6,000	6,100
Alternative 2	7,000	7,100	7,200	7,300	7,400	4,500	4,600	4,700	4,800	4,900	5,000	5,100
Savings	+2,000	+2,000	+2,000	+2,000	+2,000	−1,000	−1,000	−1,000	−1,000	−1,000	−1,000	−1,000
Year	*13*	*14*	*15*	*16*	*17*	*18*	*19*	*20*	*21*	*22*	*23*	
Alternative 1	3,800	3,900	4,000	4,100	4,200	4,300	4,400	4,500	4,600	4,700	4,800	
Alternative 2	5,200	5,300	5,400	5,500	5,600	3,300	3,400	3,500	3,600	3,700	3,800	
Savings	+1,400	+1,400	+1,400	+1,400	+1,400	−1,000	−1,000	−1,000	−1,000	−1,000	−1,000	
Year	*24*	*25*	*26*	*27*	*28*	*29*	*30*	*31*	*32*	*33*		
Alternative 1	4,900	2,600	2,700	2,800	2,900	3,000	3,100	3,200	3,300	3,400		
Alternative 2	3,900	4,000	4,100	4,200	4,300	4,400	2,100	2,200	2,300	2,400		
Savings	−1,000	+1,400	+1,400	+1,400	+1,400	+1,400	−1,000	−1,000	−1,000	−1,000		

For example, the present value of \$1 received for 5 years at 10 percent is $\mathrm{PW}(U)^5_{.10} = 3.7908$. Verify this from Table B-3 in Appendix B. Therefore, the present value of \$2,000 received annually for 5 years can be written as $2{,}000[\mathrm{PW}(U)^5_{.10}]$.

There are two steps in discounting the equal flows from period 6 through period 12. The first is to discount this uniform *stream* to the beginning of period 6, and the second is to discount this *sum* to the beginning of period 1. Perhaps a graphic illustration of this two-step process will make the logic clearer.



The present value at the beginning of year 6 for the uniform stream from period 6 through period 12 can be written as $-1000[\mathrm{PW}(U)^7_{.10}]$. Then taking this sum and finding its present value at the beginning of year 1, we have

$$-\ 1{,}000[\mathrm{PW}(U)^7_{.10}]\,[\mathrm{PW}(S)^5_{.10}]$$

where $\mathrm{PW}(S)^5_{.10}$ is the present worth factor for a sum received at the end of 5 years at 10 percent interest.

The second sequence of flows can be written as

$$\{1{,}400[\mathrm{PW}(U)^5_{.10}] - 1{,}000[\mathrm{PW}(U)^7_{.10}][\mathrm{PW}(S)^5_{.10}]\}[\mathrm{PW}(S)^{12}_{.10}]$$

The terms in the first bracket represent the discounted flows when looking from year 13 through year 24. The second bracket takes this sum and discounts it to the present period, which is 12 years earlier. The third sequence can be written as:

$$\{1{,}400[\mathrm{PW}(U)^5_{.10}] - 1{,}000[\mathrm{PW}(U)^7_{.10}][\mathrm{PW}(S)^5_{.10}]\}[\mathrm{PW}(S)^{24}_{.10}]$$

The terms in the first bracket represent the discounted flows when looking from year 25 through year 36. The second bracket takes this sum back to the present period, which is 24 years earlier.

The last two sequences are alike except for the term in the second bracket. In fact, we can write the sum for the last sequences as

$$\{1{,}400[\mathrm{PW}(U)^5_{.10}] - 1{,}000[\mathrm{PW}(U)^7_{.10}][\mathrm{PW}(S)^5_{.10}]\}[\mathrm{PW}(S)^{12}_{.10}][\,1 + \mathrm{PW}(S)^{12}_{.10}]$$

Since this pattern extends into the future, we can write

$$\{1{,}400[\mathrm{PW}(U)^5_{.10}] - 1{,}000[\mathrm{PW}(U)^7_{.10}][\mathrm{PW}(S)^5_{.01}]\}[\mathrm{PW}(S)^{12}_{.10}]$$
$$[1 + \mathrm{PW}(S)^{12}_{.10} + \mathrm{PW}(S)^{24}_{.10} + \mathrm{PW}(S)^{36}_{.10} + \cdots]$$

This represents the benefits extending from the second 12-year sequence on into the future. The only difficulty, of course, is that the evaluation of such a series extending into the future is a difficult task. However, it can be shown that a series of the type

$$1 + \mathrm{PW}(S)^h_i + \mathrm{PW}(S)^{2h}_i + \mathrm{PW}(S)^{3h}_i + \cdots + \mathrm{PW}(S)^{(n-1)h}_i$$

converges to

$$\frac{1}{1 - \mathrm{PW}(S)_i^h} \tag{3-2}$$

as n approaches infinity. That is, as the number of terms in this series increases, the sum of the series can be represented as Eq. (3-2). In our example, $h = 12$ and $i = .10$; therefore we can write our present worth of benefits, including the first 12-year sequence, as

$$\begin{aligned} \text{PW (Benefits)} = {} & 2{,}000[\mathrm{PW}(U)_{.10}^{5}] - 1{,}000[\mathrm{PW}(U)_{.10}^{7}][\mathrm{PW}(S)_{.10}^{5}] \\ & + \{1{,}400[\mathrm{PW}(U)_{.10}^{5}] - 1{,}000[\mathrm{PW}(U)_{.10}^{7}][\mathrm{PW}(S)_{.10}^{5}]\} \\ & [\mathrm{PW}(S)_{.10}^{12}]\left[\frac{1}{1 - \mathrm{PW}(S)_{.10}^{12}}\right] \end{aligned}$$

Substituting in the relevant PW factors, we have

$$\begin{aligned} \text{PW (Benefits)} &= 2{,}000(3.7908) - 1{,}000(4.8684)(.6209) \\ &\quad + [1{,}400(3.7908) - 1{,}000(4.8684)(.6209)](.3186)\left(\frac{1}{1 - .3186}\right) \\ \text{PW (Benefits)} &= \$5{,}632 \end{aligned}$$

COST STREAM

Next we turn our attention to the cost stream shown in Table 3-7. If alternative 1 is undertaken, the cost incurred will be \$10,000 less the salvage value of the defender, \$2,000, or \$8,000. Since there are no salvage values associated with present and future challengers, the cost incurred when the challenger is replaced in the years 13, 25, 37, and so on, will be \$10,000.

If alternative 2 is undertaken, the cost stream will be headed by an expendi-

Table 3-7 Stream of net costs

Year	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*
Alternative 1	8,000											
Alternative 2						10,000						
Year	*13*	*14*	*15*	*16*	*17*	*18*	*19*	*20*	*21*	*22*	*23*	*24*
Alternative 1	10,000											
Alternative 2						10,000						
Year	*25*	*26*	*27*	*28*	*29*	*30*	*31*					
Alternative 1	10,000											
Alternative 2						10,000						

ture of \$10,000 in the sixth year. Expenditures of \$10,000 will then follow in years 18, 30, 42, and so on.

The present value of the difference between these two cost streams is the figure of interest. Measuring this difference against electing alternative 1, we have

$$\begin{aligned}\text{PW (Costs)} &= \\ &8{,}000 - 10{,}000[\text{PW}(S)^{5}_{.10}] + \{10{,}000 - 10{,}000[\text{PW}(S)^{5}_{.10}]\}[\text{PW}(S)^{12}_{.10}] \\ &\qquad + \{10{,}000 - 10{,}000[\text{PW}(S)^{5}_{.10}]\}[\text{PW}(S)^{24}_{.10} + \cdots]\end{aligned}$$

As we discovered before, we can write this as

$$\begin{aligned}\text{PW (Costs)} &= 8{,}000 - 10{,}000[\text{PW}(S)^{5}_{.10}] \\ &\qquad + \{10{,}000 - 10{,}000[\text{PW}(S)^{5}_{.10}]\}[\text{PW}(S)^{12}_{.10}]\left[\frac{1}{1 - \text{PW}(S)^{12}_{.10}}\right]\end{aligned}$$

Substituting the relevant values we have

$$\begin{aligned}\text{PW (Costs)} &= 8{,}000 - 10{,}000(.6209) \\ &\qquad + [10{,}000 - 10{,}000(.6209)](.3186)\left(\frac{1}{1 - .3186}\right)\end{aligned}$$

$$\text{PW (Costs)} = \$3{,}561$$

BENEFIT/COST RATIO

It is now possible to measure the productivity of capital associated with this investment. The benefit/cost ratio can be computed as

$$R = \frac{\text{benefit}}{\text{cost}} = \frac{5{,}632}{3{,}561} = 1.58$$

It is important to realize that there is still no way to determine if the investment will be selected, for this benefit/cost ratio is placed in the pool of investment alternatives, and it is not until the final selection process takes place that it is apparent whether this benefit/cost ratio will be large enough to be among those selected. In the final analysis it boils down to the magnitude of the other benefit/cost ratios and the size of the capital budget.

A generalized form for the benefit/cost ratio The present worth of the benefit stream can be generalized to

$$\begin{aligned}&\text{PW (Benefit)} = \\ &\sum_{j=1}^{l-1} (E_D - E_C)D^j - \sum_{j=l}^{n} [(l-1)a + (l-1)b]D^j \\ &+ \left\{\sum_{j=1}^{l-1} [(n-l+1)a + (n-l+1)b]D^j \right. \\ &\qquad \left. - \sum_{j=l}^{n} [(l-1)a + (l-1)b]D^j\right\} D^n \left(\frac{1}{1 - D^n}\right)\end{aligned}$$

where l = alternative replacement year
E_D = operating expense of defender
E_C = operating expense of challenger
D = discount factor
n = economic life of challenger
a = obsolescence per year
b = deterioration per year

The present worth of the cost stream is

$$\text{PW (Cost)} = (C - S_1) - (C - S_l)D^{l-1} + C(1 - D^{l-1})\, D^n \left(\frac{1}{1 - D^n}\right)$$

where C = capital cost of challenger
S_1 = present salvage value of defender
S_l = salvage value of defender in year l ////

SUMMARY

In this chapter we looked at the problems associated with equipment replacement, while in the last chapter we focused on new investments. Although we have separated them into two chapters, equipment replacement and investments actually belong in the same category. That is, there is nothing that should make the consideration of a new investment any different from that of a replacement. True, our method for analyzing them might be different, but in the end each must stand up and face the capital-budgeting process. Only those investments or replacements which most productively employ the firm's resources should be selected.

In replacement we are usually concerned with equipment that deteriorates over time. With increasing maintenance, however, the productive life of the equipment can go on and on and on. The decision to replace is therefore not a decision to continue or not continue. Rather it is a decision to incur a capital cost so that lower operating costs will follow. Unfortunately this requires looking into the future, and Terborgh's assumptions were directed at coping with this problem. He was the first to make quite explicit the role that the future plays in making the replacement decision and what assumptions will render this problem of futurity manageable. If his model is to be used, the following question should first be answered: Are the assumptions including like adverse minima and constant inferiority gradient reasonable? If not, perhaps another model is in order. However, the additional cost incurred in developing a more complex model may be greater than the additional benefits.

QUESTIONS

Q3-1. What is the economic rationale behind the replacement of an old but perfectly functioning machine?

Q3-2. How would you design, as part of a large management information system, a replacement analysis? What data would have to be stored, what data would have to be continually gathered, and how often would you review the firm's portfolio of equipment?

Q3-3. When an old machine is replaced, what are the benefits which accrue to the firm?

Q3-4. What is the coterminated plant problem? How is this overcome by the Terborgh model?

Q3-5. What is the economic life of a project? Does it usually differ from its physical life?

Q3-6. What basic assumption is necessary in order to compute the economic life of the challenger? Is it a reasonable assumption?

Q3-7. Define the following:

(*a*) Uniform annual equivalent
(*b*) Capital recovery factor
(*c*) Obsolescence
(*d*) Deterioration

Q3-8. Theoretically, the economic life of the challenger or defender is determined by the unknown economic lives of all future machines. How does the Terborgh model resolve this dilemma?

Q3-9. In a capital-rationing situation, why is it unacceptable to simply invest in the challenger if its AM is lower than that of the defender?

Q3-10. If the total cost function shown in Fig. 3-3 is very flat in the vicinity of its minimum, what consequence might this have on replacement policy?

PROBLEMS

P3-1. Find the adverse minimum and economic life for a challenger which has the following characteristics:

Operating inferiority, $200
Capital cost, $10,000
Interest rate, 0

Use both the long and shortcut methods. By how much do they differ?

P3-2. A machine is under consideration for replacement. Its inferiority gradient is computed to be $400 per year. The capital cost of the new machine is $16,000. If the cost of capital is 15 percent, what is the challenger's adverse minimum and economic life? Find the answer by using both the long and shortcut methods.

P3-3. Solve Problem P3-1 for an interest rate of 15 percent.

P3-4. Bob Kastle cannot seem to make up his mind. Last week the new cars were unveiled, and ever since he has been unable to sleep. That new Ford LTD looks too good to believe. Reason, however, has suddenly reasserted itself, and Bob decides to undertake an economic analysis of the problem.

He is currently driving a 4-year-old Ford LTD with a value of $800. Next year it will be worth $500, the next year it will drop to $300, the following year to $200, and finally to $100, at which level it will remain indefinitely. The new LTD will cost him $3,800. Based on his experience with this model, he estimates the future resale values to be the following:

End of year	*Resale value*
1	$2,900
2	2,100
3	1,400
4	800
5	500
6	300
7	200
8 and on	100

He estimates that the maintenance costs associated with his 4-year-old car would run approximately $350 next year. The maintenance on the new car, however, would be $50.

If the opportunity cost of money to Bob is 10 percent, what action should he take?

P3-5. Plot a total cost function for a popular American automobile. Is the function relatively flat in the region of its minimum? If so, how should this influence your strategy as a car owner? Would you expect the function to look any different for an imported car?

Hint: To find resale values, use either the newspaper or blue-book values. Consider all costs which vary with auto value. These might include insurance, registration fees, and excise taxes.

//// **P3-6.** The economic life of a particular challenger has been calculated to be 10 years. It accumulates obsolescence at $200 per year and deterioration at $200 per year. For the first year of the comparison period, the challenger is expected to have an operating cost of $8,000 while the best available defender will have an operating cost of $10,000. The capital cost of the challenger is $20,000 and it has no salvage value at the end of its economic life. The salvage value of the defender, on the other hand, is $5,000 in the beginning of the first year, $2,000 in the second year, and zero in the third. The economic life of the defender is computed to be 2 years. Using an interest rate of 10 percent, calculate the benefit-cost ratio associated with purchasing the challenger now rather than waiting for 2 years. ////

P3-7. You are considering the replacement of a 10-year-old machine with one which is numerically controlled. The new machine costs $60,000 and will save you the following:

Repairs	$3,000
Downtime	5,000
Rejects	3,000
Power	1,000

The defender has a salvage value of $10,000 now, but that will decline to $6,000 next year and to $4,000 the next; it will finally remain at $2,000 each year thereafter. Should the new machine be purchased if the firm's cost of capital is 12 percent?

P3-8. A machine shop is considering the purchase of a new grinder to replace a 15-year-old machine of the same kind. The new machine will cost $15,000. The present machine, on the other hand, is worth $7,000 now and will be worth $6,000 in one year.

There are several advantages associated with the new machine. First there will be a direct labor savings of $3,100 per year. Then, because the machine produces a product of high quality, there will also be a $1,000-per-year direct labor savings on subsequent operations. The maintenance that will be saved next year is estimated to be $300; and since the machine takes less floor space than its predecessor, there will be an additional $100-per-year saving.

Using the shortcut method to determine the AM of the challenger and defender, choose a course of action for a 10 percent cost of capital.

P3-9. The city of Bridgeport is considering the purchase of a new snowplow to replace a 10-year-old unit. The new machine, however, is much more versatile and also more productive. It can accomplish 20 percent more work in the same period of time. This would mean that the city would save $5,000 in payments to contractors who plow snow for the city. In addition, the new unit will save about $1,000 next year in maintenance costs.

The cost of the new unit is $14,700 and the resale value of the old unit is $1,000 now and will drop to $500 next year. If the city's cost of capital is 10 percent, should they make the trade?

P3-10. Solve Prob. P3-4 when the opportunity cost of money is 8 percent. How sensitive is the decision to the opportunity cost of money?

P3-11. A large manufacturing firm is considering the replacement of a machine tool with one which is numerically controlled. The new machine will cost $100,000. After careful consideration they decide that the assumption of constant inferiority accumulation is inadequate. Instead they predict that the inferiority gradient will be as follows.

Year	*Inferiority*
0	$ 0
1	1,000
2	3,000
3	6,000
4	10,000
5	15,000
6	21,000
7	28,000
8	36,000
9	45,000
10	55,000
11	55,000
12	55,000
13	55,000
14	55,000
15	55,000

If the cost of capital is zero and there is no salvage value, what is the economic life of this challenger?

P3-12. If the cost of capital is 10 percent in Prob. 3-11, what is the economic life of the challenger?

BIBLIOGRAPHY

Brennan, J. F.: "Optimum Life of Fleet Automobiles," *Journal of Industrial Engineering*, November–December 1964.

Gillis, Floyd E., Jr.: *Managerial Economics*, Addison-Wesley, Reading, Mass., 1969, chap. 5.

Grant, Eugene L., and W. Grant Ireson: *Principles of Engineering Economy*, The Ronald Press Company, New York, 1970.

Mayer, Raymond R.: "Problems in the Application of Replacement Theory," *Management Science*, April 1960.

Meyer, Robert A., Jr.: "Equipment Replacement under Uncertainty," *Management Science*, vol. 17, no. 11, July 1971, p. 750.

Smith, Vernon L.: "Economic Equipment Policies: An Evaluation," *Management Science*, October 1957.

Terborgh, George: *Dynamic Equipment Policy*, Machinery and Allied Products Institute, Washington, D.C., 1949.

4

Process Design: Macro

INTRODUCTION

A *transformation process* can be defined as a set of operations that are performed at stations for the purpose of transforming inputs into outputs. For example a TV manufacturer transforms material inputs into a TV set, a hospital transforms an ill patient into a healthy one, a school transforms a nonreader into a reader, a service station transforms a sluggish automobile into a well-running machine, and a loan department at a commercial bank transforms a loan request into either a loan or a rejection. In each of these situations an input was processed through the system and an output was generated. The processing systems were, of course, quite different, but nonetheless we will see that there are conceptual similarities that exist between them.

The design of a transformation process requires that several questions be asked: What are the characteristics of the product or service being supplied? What is the expected volume of output? What kinds of equipment are available, or must the equipment be custom built? What is their cost? What types of labor skills are available, in what quantities, and at what wage rate? How much money can be spent on the process? Should the process be labor- or capital-intensive? Should the component parts be made or purchased? And if material handling will be a problem, how will it be solved? These and countless other questions must be asked and answered before the design of a facility is finalized.

This chapter focuses on some of these issues. First a framework will be developed under which the topics of process design can be placed. Then we will turn to the first problem, that of work station selection. Here the decision is made as to whether the process will be capital- or labor-intensive. Next we will examine the usefulness of schematic aids in the design process. Then the focus shifts to work flow considerations and finally to the make or buy decision.

A FRAMEWORK FOR PROCESS DESIGN

Process design can be viewed as an iterative exercise. That is, problems are solved one at a time and sequentially; then after each stage, or perhaps after several stages, the previous stages are reexamined to see if later steps have affected the best way in which these steps should have been designed. This procedure is illustrated in Fig. 4-1.

PRODUCT PLANNING

Product planning serves as an input to process design. However, in most cases the responsibility for this phase rests with groups, such as marketing and engineering, which are generally found outside the domain of process design. It is early in this stage that the perceived needs of the consumer are identified. If the process is service-oriented, these needs will be reflected in the proposed quality, speed, cost, and reliability of the service. If, on the other hand, the process will be manufacturing-oriented, then these needs will be reflected in the product's proposed quality, cost, function, reliability, and appearance. When these service or product parameters are transformed into a product design, it is essential that a cross-functional alliance between product planning and process design groups be established in order to ensure that the product objectives can indeed be profitably met. Otherwise only local goals may be pursued. For example, the marketing and engineering departments working alone may design a product which is very costly to manufacture and very difficult to service. In short, the transformation process should be considered well before the design is finalized. This can usually be accomplished in the following way.

Information from the *product development* stage can be directed to those responsible for *process R&D*. They in turn can determine if the process capability for this product now exists within the firm, whether it exists outside the firm, or whether research and development effort would be necessary to meet the specified objectives. Rough cost estimates would also be made for each alternative identified. In large firms a special process R&D department may be organized just for this purpose. In smaller firms this function is less formalized and more reliance would be placed on outside suppliers of processing equipment and company engineers.

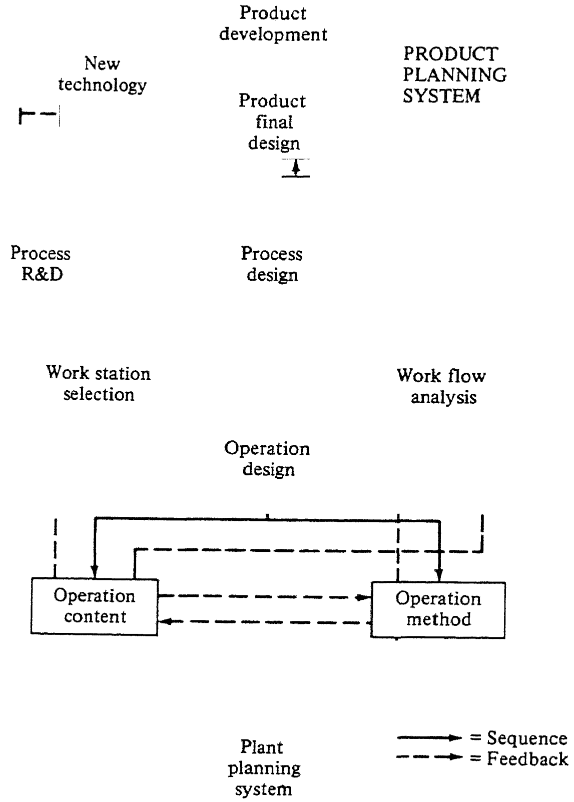


Fig. 4-1 The process planning task. (*From Howard L. Timms and Michael F. Pohlen, The Production Function in Business, Richard D. Irwin, Inc., Homewood, Ill.*, 1970, *p*. 302.)

As the process R&D phase progresses, information is fed back to the product development group. If the R&D group will be able to comply with the product specifications, then the *product final design* stage can begin. If there are problems, however, in complying with product specifications, then modifications in the product development stage must occur.

Information from both product final design and process R&D are inputs to the *process design stage*. The purpose of this stage is to generate alternative ways of meeting the objectives formalized in the final design stage, determine the criteria by which they will be evaluated, and make the final selection.

PROCESS DESIGN: MACRO

Process design: macro is composed of two aspects: the choice of work station and the choice of work flow. Work station selection involves the choice of machines to be included in the process, whereas work flow analysis concerns the flow of work between these stations.

It is here that the decision is made as to whether the process will be continuous, intermittent, or some combination of both. In *continuous processing,* the process is in constant operation and usually involves a high capital-to-labor ratio. Typical of a continuous-process orientation are the automobile industry, chemical processing, plastics, some high-volume electronic manufacturers, and some utilities such as telephone, power, and gas transmission. Also characteristic of continuous-process industries is a *product layout* where all the work stations are devoted exclusively to a single product and are grouped according to the processing requirements of that product. In more cases than not, the machines found at these work stations are special-purpose and costly and have little versatility outside their own product line. Work flow is largely specified by the physical characteristics that the product layout takes. Often a conveyor system is used. An example of this kind of process is an assembly line for machining engine blocks at an automotive plant.

At the other end of the spectrum we find *intermittent processing.* Here production or service for any one job is carried on intermittently, not continuously. There is a high mix of products which use the facility, and portions of the process may be in operation several times during the day or only occasionally during the month. Usually the relatively low demand for each product or service does not warrant the high investment in a continuous process. Typical of intermittent processes are job shops, emergency rooms in hospitals, hospital laboratories, most office work, many educational processes, and most services. The predominant plant layout that one expects in intermittent processing is a *process layout.* In a process layout, machines or services of the same category are grouped together. We therefore find lathes, milling machines, inspection stations, and so on, in one location.

A characteristic of the process layout is that the particular sequence of operations that one job follows is seldom repeated by other jobs flowing through the process. For example the sequence from admittance to discharge in a hospital is seldom the same from one patient to the next. A patient may or may not have X-rays, may or may not have surgery, may or may not undergo physical therapy, etc. Consequently the work flow is not nearly as predictable as in the continuous-process case, and we therefore find variable-path equipment, such as forklift trucks, employed in processes of this type.

The criteria used to make the choice between these two layouts include investment costs, material handling costs, direct labor costs, space requirements, equipment flexibility in meeting changes in output quantity, system reliability, and maintenance costs. These economic and noneconomic factors must, in turn, be weighed before making the final decision.

PROCESS DESIGN: MICRO

In the next stage attention shifts to the details that make up the work at each station. Concern is with the operational content and operational method of the task. *Operational content* focuses on the appropriate combination of steps that should be assigned to a work station. *Operational method,* on the other hand, is concerned with the efficient execution of these steps.

PLANT-PLANNING SYSTEM

If a new plant will be built, then process planning proceeds in a relatively constraint-free setting. If, however, the process must be installed in existing plant, then a set of spacial constraints must be observed. Therefore the *process-planning system* must interact with the *plant-planning system* to ensure feasible layouts.

PROCESS DESIGN AS AN ONGOING ACTIVITY

Process design is not strictly limited to new design. Whenever the costs of inputs change by a substantial amount, or whenever output levels or quality objectives change, process review should be initiated. If the reason for the reexamination is a price change in an input factor, then a new input mix reflecting the different price ratio of the substitute inputs may be desirable. For example, if wage rates increase substantially as a result of a new labor contract, it might be possible that automated sequences in the production line which were inefficient uses of capital before might now be profitable.

Process design, then, should be a continuous activity—not one which is precipitated only when new products or services are introduced. Since it is continuous, the question of how much money to spend in its pursuit is of utmost importance. The amount which should be spent depends, of course, upon the benefits accrued. At the limit, effort should continue as long as the marginal benefits from the improvement are greater than the marginal costs. Again we use the word *marginal,* since it is only those costs that change which are relevant. Fixed costs and some semivariable costs, to the extent that they are not affected by the redesign, are irrelevant.

For firms at the continuous-process end of the spectrum, considerable effort can be channeled in this direction. Small improvements made in the process are magnified when the volume going through each stage is large. Cost savings in the automotive industry, for example, are measured in fractions of a cent per unit. In the service industries, on the other hand, savings are usually measured in dollars, but the volume through each station is, of course, much less.

If redesign requires investment, it seems reasonable to subject the investment to the review of the capital-budgeting process. Therefore this investment, along with all the others which the firm is considering, must face the final selection process in which only the most profitable survive.

Case study: The process design of progressive patient care facilities The administrator for the Smithdale Hospital, Ken Swenson, has been quite concerned about the low level of occupancy and the inefficient service which have characterized several of his new facilities. Five years ago an intensive care zone was added, three years ago a coronary care zone was developed, and last year a self-care zone was added for patients who were beyond the need for the traditional intermediate care but were not well enough to be discharged.

When the initial design was undertaken for each of these zones, they were considered as independent entities. What resulted was a group of small, exclusive, independent services which had a low occupancy rate and which utilized staff rather inefficiently.

In an effort to improve the efficiency of the zones, Mr. Swenson decided to take a fresh design approach. Rather than treat each section independently, he would take a systems point of view. He received his inspiration from a new concept called *progressive patient care.*

In progressive care, the patient is initially admitted to a zone determined by his health status. Each zone has predetermined medical criteria with which this status must be matched. For the Smithdale Hospital, Mr. Swenson has divided the zones in the following way:

1. Intensive care
2. Coronary care
3. Intermediate care
4. Continuing care
5. Self-care
6. Psychiatric care

A central aspect of this concept is that the patient may be moved from one zone to the next according to his health status. From a design point of view, this means that when the capacity of the system is determined, the zones must be looked at, not singly, but as components of a system.

The design problem is to determine the size of each zone that best meets minimum service levels while not violating the available operating and investment funds.

Mr. Swenson decided to call for help at this stage. He felt that a model could be built which would determine the fraction of his beds that should be assigned to each zone. After discussing this problem with several management consultants, he signed a contract with Crosby, Stills and Nash, Inc.

Their approach was to simulate the consequence of several different strategies. In each of these strategies a different fraction of the total number of hospital beds was assigned to each zone. The simulation proceeded by processing patients through the hospital, first admitting them to the appropriate zone and then transferring them to other zones as their health status changed or discharging them from the hospital. The data that the consultants used were taken from the hospital records.

For each strategy, the output from the simulation summarized the percent of occupancy and the number of times that service levels were violated. From this output, Mr. Swenson chose what he thought was the best strategy and began to prepare his case for presentation at the next trustees' meeting.

WORK STATION SELECTION

After the product final-design stage has been completed, the selection of the equipment to be included at the work station must be made. In the usual case there is more than one way in which a part or a demand for service can be processed through the system. For example, if a hole has to be threaded, a turret lathe, engine lathe, or numerically controlled machine may be used. In the hospital laboratory, the blood samples can be analyzed by a technician or they can be analyzed by a machine. In the classroom the student can learn French by classroom instruction or by videotaped lectures or through programmed instruction.

SELECTION OF EQUIPMENT BY ECONOMIC COMPARISON

Suppose that the criterion for process design is cost minimization. If the necessary equipment for these alternatives is already in use and if sufficient capacity exists to meet this new demand, the choice will depend upon the one which can process the forecasted volume at the lowest operating cost. The total operating cost of processing a job on any alternative can be broken down into its fixed and variable components. Fixed costs include setup and tooling charges, whereas variable costs include material, labor, and power. Given several alternatives, then, the most economical will have the lowest sum of these fixed and variable costs at the expected volume.

Consider the following example. A part must be processed through one of three possible machines. These include an engine lathe, a turret lathe, and a numerically controlled machine. The fixed and variable costs associated with each machine are presented in Table 4-1. A breakeven chart is drawn in Fig. 4-2, where the total costs over a range of outputs for each alternative are shown. It can be seen that if volume is less than V_1, alternative 1 is the best; if volume is between V_1 and V_2, then alternative 2 is the best; whereas a volume greater than V_2 should lead to the choice of alternative 3.

This simple least-cost comparison assumes that these machines are currently owned and operating and that capacity exists to accommodate this increased usage. If, on the other hand, the machines must be purchased, the

Table 4-1 Fixed and variable costs for three machine alternatives

Machine	*Fixed costs*	*Variable costs*
1. Engine lathe	$10	$0.30
2. Turret lathe	40	0.15
3. Numerically controlled machine	75	0.03

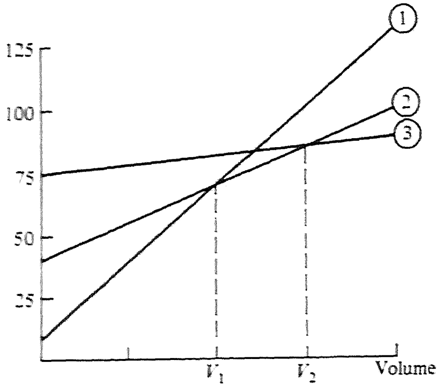


Fig. 4-2 Breakeven chart for three machine alternatives.

breakeven analysis is irrelevant and an investment analysis, perhaps focusing on the benefit/cost ratio, must be undertaken.

NONECONOMIC FACTORS

Selection of the final process is many times not based on economic criteria alone. Quality, speed, and precision represent other bona fide criteria. If, for example, cost is the only criterion, the ranking of three alternative machines might be 3, 1, 2. That is, if machine 3 is used, the part will be produced at lowest cost. Next is machine 1 and then machine 2. If speed of manufacture was of utmost importance so that a delivery date would be met, the priority might be 1, 3, 2. Therefore the priority depends upon the criterion. In some cases the criterion for the same product or service from one run to the next might not be the same. For example, in a routine production run, the criterion for selecting a process might be strictly cost. If, on another occasion, this part is holding up a major shipment, the criterion becomes time, and machines which were out of the question during the previous run are just what is needed on this run. In a hospital environment, routine X-rays might be taken in Unit 1. This would ensure the lowest cost per X-ray. In emergencies, however, if there is a waiting time at Unit 1, then Unit 2, usually reserved for more specialized cases, might be used.

Therefore, when multiple criteria are likely to apply, it might be quite reasonable to specify more than one alternative on the routing sheet. This type of flexibility can be uncovered in job shops, service-oriented industries, hospitals, and generally in the intermittent-process industries. The decision as to what alternative to take might be left until a particular job is dispatched to the shop. At that time the relevant criterion might be applied.

SCHEMATIC AIDS IN PROCESS DESIGN

In the planning of a complex process, it is often quite helpful to use schematic charts. It is through these charts that the logical sequence of a process can be examined and improved.

During the last stages of product final design, the engineering process being nearly complete, assembly drawings which show the relationship of the parts to the whole should be made available to those responsible for process planning. These generally serve as the starting point for the generation of one or more processing plans.

In developing such a plan several steps are followed. First the set of operations which must be performed is determined, followed by the sequence in which they must be executed. Then the equipment which is necessary is identified, as is the specification of special tools and fixtures. Sometimes the estimate of setup and run times is also made.

OPERATION PROCESS CHART

A frequently used chart for illustrating this information is the *operation process chart*. Two basic symbols are used, one for operations and the other for inspections. In Fig. 4-3 an operation process chart is shown which illustrates the steps necessary for the manufacture of a bracket.

When the operation process chart is completed, it can be used to question whether or not certain operations and inspections are necessary and to determine if some rearranging could improve the physical flow.

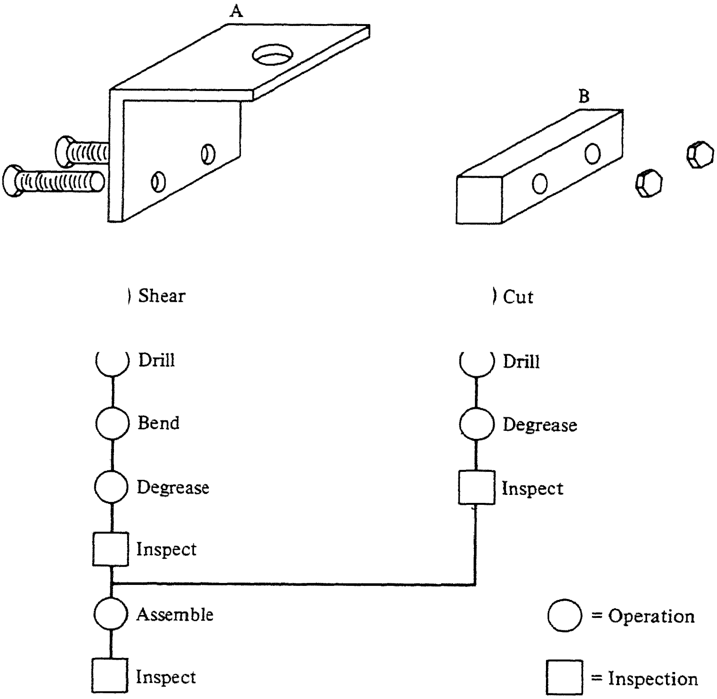


Fig. 4-3 Operation process chart.

FLOW PROCESS CHART

A more complete analysis of physical flow can be determined by the incorporation of nonproductive activities, such as storage and transportation, into the operation process chart. With these added dimensions we then have a *flow process chart,* shown in Fig. 4-4. Since these additions are nonproductive, it is in the best interest to try to minimize them. Only in totally automated processes will it be possible, however, to nearly eliminate these unproductive stages. As the process becomes more labor-intensive and the orientation goes from continuous process to intermittent process, the frequency of transportation and delays will increase. In the intermittent processes, temporary storage, however, does sometimes serve an important function. It decouples one stage from the next. Only in this way can production proceed smoothly, for without these inventories work stations might remain idle for long periods of time before they are supplied with work. There will be more on this when we discuss line balancing in the next chapter.

In continuous-process design, the process designer should develop several alternative flow process charts before making a choice. For each alternative, every step should be questioned; nothing should be taken for granted. Can steps be eliminated? Can steps be combined with other steps? Does inspection take place at the right stage in the process? Can the sequence be altered? Are transportation and storage being accomplished in the right way? With this kind of approach, a reasonably good design is assured.

Once a number of alternative processes have been developed, the next problem is to evaluate them and choose the best one. In continuous-process industries, for example, the criterion is generally cost minimization. Therefore a summation of all costs for each alternative is compared. The total cost for alternative j is

$$C_j = \sum_{i=1}^{n} (T_{ij} + S_{ij}) + M_j + T_j$$

where C_j = total cost for alternative j
T_{ij} = transportation costs to work station i
S_{ij} = cost of the operation performed at work station i
M_j = raw material cost for process j
T_j = transportation costs from last work station to inventory in process j
n = number of work stations in process j

To establish the cost for each operation, standard times are frequently used. To determine a standard time, one approach is to break the job down into its basic elements. The times for these basic elements can be found in prepared manuals. Then these times are added up to determine the total time for the operation. In the next chapter the details behind standard times will be examined.

When all the alternatives have been specified and the costing process is

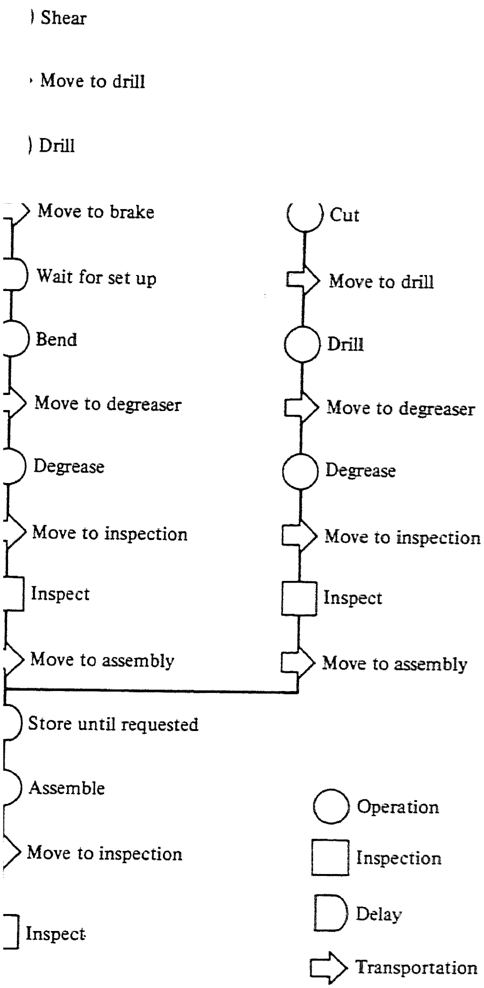


Fig. 4-4 Product flow process chart.

complete, the selection is made. At this time many imponderables must be considered. Can the process be expanded if demand should increase? Is it flexible enough to accommodate changes in product design? Is there some operating flexibility, so that the machines might be used to manufacture other products? Is the process capable of variable output levels? All of these factors must be considered before the final decision is made and a process irrevocably selected.

For intermittent processes, the selection of the process or processes is not nearly as irrevocable. The equipment is usually of the general-purpose variety, and it is likely to be used for many other jobs. In fact, there are often several alternative routings which a product can take through the process. They could all be identified on the routing instructions; and when the job is dispatched, its appropriate route can be chosen in full knowledge of the existing loads on the process—knowledge which it is impossible to have in the planning stages.

WORK FLOW: MATERIAL HANDLING

The transportation symbol is usually quite prevalent in many flow process charts. Therefore the transportation system can markedly affect both manufacturing costs and production time. The choice of materials handling equipment, then, should play an important role in process design.

In the two types of flow, continuous and intermittent, two very different types of material handling systems are used. In the continuous process are found fixed-path systems. These include overhead conveyors, roller conveyors, etc., whose path between work stations in the process is fixed. Variable-path equipment, on the other hand, is free to travel in any direction and to any work station. Typical of variable-path equipment are forklift trucks and overhead cranes. They are common to intermittent processes.

The choice of the specific piece of equipment depends upon several factors, including the physical characteristics of the part to be handled, the quantity to be transported, the speed, and the spatial configuration of the building. If the handling system must hold the work during an operation, this becomes another dimension which should be considered in the selection process.

Once the type of equipment has been identified, the purchase must be justified by capital investment analysis. The purchase of materials handling equipment is like the consideration of any other piece of capital equipment: it must face other alternative investments in the competition for scarce funds.

Case study: Material handling The Grandiron Company is a job shop which specializes in the machining of large sections of sheet metal and large castings. Business has been good and getting better, and as a result the backlog has been increasing. The production manager, however, attributes some of this problem to an inefficient materials handling system.

At present the materials handling system consists of one forklift truck which carries the work from raw material storage to the first work station, between work stations, and from the final work station to finished goods inventory.

When a machine operator needs service, he notifies the tool crib attendant, who in turn communicates with the forklift operator by mobile radio. If there are several calls for service in a short interval, this service is extended on a first-come, first-served basis.

Recently there has generally been a queue waiting for service. The consequence of this is lost time, over which the machinist has little control. If he is waiting for material or waiting for the finished job to be taken from his station to the next station, there is very little work which he can do.

The production manager therefore feels that the current materials handling system is partially to blame for the large backlog. He feels quite certain that if a more effective system were used, the backlog would be reduced.

To analyze this problem in more detail, the production manager brings in the industrial engineering department. He asks them to study the problem and suggest a solution. Their first step is to identify alternative material handling methods. They decide to evaluate two alternatives. The first is the addition of the second forklift truck to the shop floor. The second is to retain the status quo.

From these two alternatives it will then be possible to evaluate the economic advantage that the addition of the second truck has over the operation of a single truck. This advantage is in the form of a stream of cost savings extending over the life of the second forklift truck. The present value of this stream represents the benefits, and the ratio of this sum to the capital cost of the truck—the benefit/cost ratio—represents the relative attractiveness of the investment.

How is the savings stream determined? First the cost of waiting must be estimated. What is the opportunity cost of an idle minute spent while waiting for forklift truck service? After much consideration the industrial engineer sets this waiting-time cost at 30 cents per minute. Next the performance of the single truck must be established by simulation, and from this the yearly waiting costs can be estimated. This is followed by a simulation of the two trucks working simultaneously. The yearly waiting cost associated with the two trucks should be lower than that with the single truck, and the magnitude of this savings, extended over the project's life, is used to determine the present value of the benefit stream. The difference in costs between the alternatives determines the relevant cost in the benefit/cost ratio.

The simulation requires that the frequency distribution for time between arrivals and the distribution for the time it takes to complete a service call be determined by sampling. One hundred observations were taken, and the outcome can be seen in Table 4-2. Also added to the table are the relevant two-digit numbers to be used in the simulation process. First a simulation of the present situation is undertaken. The sequence of steps in carrying out this simulation is summarized below.

1. Generate a two-digit random number.
2. Identify the arrival time associated with the random number.
3. Generate another two-digit random number.
4. Identify the service time associated with the random number.
5. Determine the earliest time that service can begin.
6. Determine when service will end.
7. Compute waiting time by subtracting the time that service starts from the arrival time.
8. Repeat steps 1–7.

Table 4-2 Time between arrivals and service time for 100 arrivals

Time between arrivals

Time	*Number of occurrences*	*Two-digit numbers*
3	5	00–04
4	10	05–14
5	14	15–28
6	27	29–55
7	31	56–86
8	10	87–96
9	3	97–99
	100	

Service time

Time	*Number of occurrences*	*Two-digit numbers*
4	10	00–09
5	25	10–34
6	35	35–69
7	20	70–89
8	10	90–99
	100	

We start by selecting a number from a random number table. It can be seen that 17 was drawn first. Referring to Table 4-3, the time between arrival of a call and service associated with this number is 5 minutes. If the shop begins work at 8:00 A.M., the first call for service occurs at 8:05. Since there is no one waiting, service can begin immediately. Generation of the second random number, 22, reveals that the service time will be 5 minutes and that service will be complete at 8:10. There is no time spent waiting.

The second arrival occurs 3 minutes after the first, at 8:08; however, service is not available until 8:10, at which time the truck completes its first call. Therefore waiting time for the second call is 2 minutes.

This simulation is continued for two hours. In these two hours we can see that 53 minutes of waiting time occurred. The waiting time for the day could be approximated as $53 \times 4 = 212$; and if there are 250 working days in a year, the waiting time in one year can be approximated as $212 \times 250 = 53{,}000$ minutes. At a waiting cost of 30 cents per minute, the total waiting cost per year with one truck is $53{,}000 \times \$.30 = \$15{,}900$.

Now we turn to the analysis with two trucks. The only difference

Table 4-3 A two-hour simulation of the present situation

Random number	*Time between arrival*	*Clock time at arrival*	*Service starts*	*Random number*	*Service time*	*Service ends*	*Waiting time*
17	5	8:05	8:05	22	5	8:10	0
02	3	8:08	8:10	50	6	8:16	2
64	7	8:15	8:16	13	5	8:21	1
97	9	8:24	8:24	36	6	8:30	0
77	7	8:31	8:31	91	8	8:39	0
85	7	8:38	8:39	58	6	8:45	1
39	6	8:44	8:45	45	6	8:51	1
47	6	8:50	8:51	43	6	8:57	1
09	4	8:54	8:57	36	6	9:03	3
44	6	9:00	9:03	46	6	9:09	3
33	6	9:06	9:09	46	6	9:15	3
01	3	9:09	9:15	70	7	9:22	6
10	4	9:13	9:22	32	5	9:27	9
93	8	9:21	9:27	12	5	9:32	6
68	7	9:28	9:32	40	6	9:38	4
27	5	9:33	9:38	68	6	9:44	5
66	7	9:40	9:44	26	5	9:49	4
73	7	9:47	9:49	85	7	9:56	2
70	7	9:54	9:56	11	5	10:01	2
62	7	10:01					53

in this simulation is that whenever a service call is received, the first available truck is dispatched to the scene. Since there are two trucks, we would expect waiting time to be sharply reduced.

In Table 4-4 a 2-hour period is simulated. Upon arrival of the first call, truck 1 is dispatched. It completes service at 8:14. Call two arrives at 8:12 and is assigned to truck 2, since truck 1 is still servicing the first call. Call three arrives at 8:20 and both trucks are available. It does not matter which truck is assigned.

The simulation is continued until 10:00, and it can be seen from Table 4-4 that no waiting occurs. Therefore there are no waiting costs associated with the two-truck simulation.

Next the difference in cost between these two alternatives must be determined. Only those costs which are not common to both alternatives need be considered. For example, the first truck and its operator will be kept in both alternatives; therefore these costs need not enter into the analysis.

In the second alternative, a truck which costs $15,000 and an operator at a yearly cost of $12,000, including fringe benefits, will be required. If the life expectancy of the new truck is 10 years, the savings generated with the two trucks is equal to the savings associated with the reduced waiting time less the cost of the second operator: $15,900 — $12,000 =

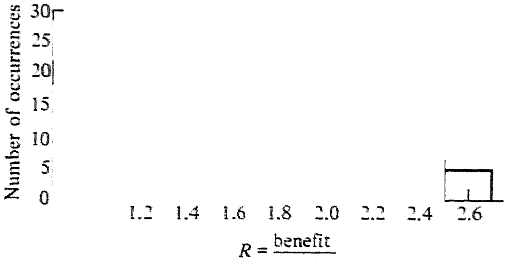


Fig. 4-5 Benefit/cost distribution.

\$3,900. This is a savings which occurs each year for 10 years, and from the annuity table the present value of this stream at a 10 percent cost of capital is 3,900(6.145) = \$23,965. The benefit/cost ratio is therefore

$$R = \frac{23,965}{15,000} = 1.6$$

A single simulation yields but one benefit/cost ratio, whereas another simulation would yield another ratio. Therefore several replications must be undertaken, after which time a distribution of benefit/cost ratios can be constructed. Such a distribution for 100 replications is shown in Fig. 4-5. The average benefit/cost ratio is 1.8, and from the graph a measure of the riskiness of this investment can be estimated by the dispersion of the benefit/cost distribution.

THE MAKE OR BUY DECISION

It is in the process design stage that major make or buy decisions must be made. The choice is between producing the component part within the firm and purchasing it from an outside supplier. Nearly every component in the product is a candidate for this decision and therefore should be carefully scrutinized on both economic and noneconomic grounds. First we take a look at economic considerations.

Consider the problem in determining the relevant cost considerations when a component is bought instead of being made. Only those costs that are directly affected by the decision need be analyzed. For example, if supervision requirements, floor space, heat, and light are unaffected by a decision to buy, then none of these costs should occur in the economic analysis. Therefore one must ask, What costs will be avoided if the component is bought instead of being made? If, on the other hand, the decision is to make instead of buy, then the relevant costs are those which will be incurred by the addition of this component to the manufacturing load. The following example illustrates these economic problems and some noneconomic problems involved in the make or buy decision.

Table 4-4 A two-hour simulation with two trucks

					Truck 1		*Truck 2*		
Random number	*Time between arrivals*	*Clock time at arrival*	*Random number*	*Service time*	*Service starts*	*Service ends*	*Service starts*	*Service ends*	*Waiting time*
65	7	8:07	70	7	8:07	8:14			
23	5	8:12	66	6			8:12	8:18	
90	8	8:20	92	8	8:20	8:28			
78	7	8:27	79	7			8:27	8:34	
70	7	8:34	88	7	8:34	8:41			
85	7	8:41	90	8	8:41	8:49			
97	9	8:50	54	6	8:50	8:56			
84	7	8:57	12	5	8:57	9:02			
20	5	9:02	10	5	9:02	9:07			
05	4	9:06	02	4			9:06	9:10	
35	6	9:12	01	4	9:12	9:16			
37	6	9:18	51	6	9:18	9:24			
94	8	9:26	17	5	9:26	9:31			
00	3	9:29	52	6			9:29	9:35	
77	7	9:36	40	6	9:36	9:42			
80	7	9:43	75	7	9:43	9:50			
32	6	9:49	33	5			9:49	9:54	
69	7	9:56	80	7	9:56	10:03			
									0

Case study: Make or buy The Sanbourne Company, a manufacturer of electronic instrumentation, is currently purchasing several toroidal transformers. These transformers are used in many of their instruments and their use will probably increase in the future. In the manufacture of these transformers, a coil-winding machine winds an insulated wire around a doughnut-shaped core which is approximately $\frac{1}{2}$ to 2 inches in diameter, depending upon the particular toroid being wound. Up to now all of the transformers have been purchased from a vendor. Their prices have seemed reasonable and the quality acceptable; in addition, they have supplied valuable engineering assistance on several occasions.

To study the feasibility of starting a toroidal manufacturing facility in the Sanbourne Company, the production manager chooses to focus on the unit that is purchased in the largest quantity. The purchase cost for this unit is $4.10, and last year 5,000 of them were used.

Although the company at present does not have any capability for mass-producing transformers, they do have, as part of the research lab, a coil-winding machine and an experienced operator. This facility, however, is limited to the manufacture of one or two at a time and is primarily a service to engineers who are experimenting with different circuit designs.

A discussion between the production manager and the supervisor of the R&D shop leads to the conclusion that there is enough capacity on the coil-winding machines in the R&D shop to wind the 5,000 toroids and perhaps even more. The marginal costs of making rather than buying will therefore include only labor and material costs. No other costs are expected to increase for these 5,000 toroids. Totaling these up, the expected marginal manufacturing cost is $1.95 each. Economically this certainly seems to be a reasonable venture.

When it comes time to bring the rest of the toroidal transformers into the plant's manufacturing process, new coil-winding equipment will have to be purchased. At that time, it will no longer be a simple marginal cost analysis, since a substantial investment must be undertaken. It will become a capital investment problem. The cost and benefit streams must be estimated, a benefit/cost ratio determined, and the project subjected to the capital-budgeting selection process.

For the time being, the only relevant costs are labor and materials, and economically it appears to be a sound decision to manufacture these 5,000 units. There are, however, several noneconomic factors that normally play a role in a decision such as this. It was indicated that the Sanbourne Company expected the demand for these transformers to increase. Therefore it is quite reasonable to expect that they would like to gain some expertise in the manufacturing technology behind these components. This could perhaps give them a better chance to design the units for more economical manufacture. They might even indulge in some process R&D. And, as some experience with the manufacturing process is gained, manufacturing costs might even drop below current expected levels.

> Furthermore, once these improved manufacturing techniques are operative and successful, they become proprietary information and Sanbourne would hold a competitive advantage. This would not be true if they bought the units from a vendor, for any improvements the vendor made would be passed on to his customers.

There are, of course, factors other than those illustrated in the case which are frequently taken into consideration during the make or buy decision process. A very important one has to do with the control of quality. If the component is purchased, in most cases it is only when delivery is made that its quality can be observed. Even then, the only decision that can be made is to accept or reject the lot.

Another benefit of making rather than buying is that the availability of supply is assured. Reliance does not have to be placed on an outside source which is beyond the reach of the firm's control. When the item is made, additional resources can be employed if delivery must be speeded or if bottlenecks occur. This strategy is unavailable if the items are bought.

In some circumstances the firm has no alternative but to manufacture. This occurs when the manufacturing technology is unavailable elsewhere or the process is protected by patents.

The level at which the process is operating affects marginal costs, and therefore it is quite likely that a decision to make during one quarter may be followed by a decision to buy in the next. As capacity is approached, marginal costs run higher than if the process was operating at less than full capacity. For example, in a period of heavy demand, the marginal costs associated with a component might include an overtime premium, a second shift, the rental of additional floor space, and so on. This cost might be so high that a more economical strategy would be to subcontract some or all of the order quantity and resume manufacturing when the shop load is reduced.

When shop loads play a role in the make or buy decision, and some components which were made in one period are purchased in another, the domain of production scheduling has been entered. For by changing our make or buy strategy in response to shop load, we have really utilized a method for absorbing fluctuations in demand. It is, of course, the role of production scheduling to decide how best to absorb fluctuations in demand. We wait until Chap. 13 for a more detailed analysis of aggregate production planning and how the make or buy decision can play a role in achieving an economical schedule.

SUMMARY

In this chapter we considered some major problems which must be solved during the first stages of process design. It must be determined whether or not the process will be continuous or intermittent, what machines will be employed, and what will be made and what purchased.

The design of a processing system is still more art than science. Nothing yet developed can replace the creativity that a process designer can bring into the design stage. There are some guidelines and some techniques that are helpful, but it is those imponderables and special situations that continue to make process design itself a labor-intensive activity.

QUESTIONS

Q4-1. What is a transformation process, how can they be categorized, and what is the predominant layout pattern for each of these categories?

Q4-2. What feedbacks are essential in the process design stage?

Q4-3. In the progressive patient-care case, suppose that the choice of a system design has been narrowed down to two alternatives. The results of the simulation analysis show the violation of service levels to be the following:

	Alternative 1	*Alternative 2*
(*a*) Intensive care	3%	2%
(*b*) Coronary care	1	4
(*c*) Intermediate care	5	4
(*d*) Continuing care	6	4
(*e*) Self-care	9	11
(*f*) Psychiatric care	11	8

Which one would you choose? What was your criterion? Why is the choice difficult? Suppose that a violation of the service level simply means the request for a bed in a particular zone when none is available. In what ways could the processing of patients through the system be modified to avoid some of these violations?

Q4-4. What are the differences between operation process and flow process charts?

Q4-5. What are the relevant costs in a make or buy decision?

Q4-6. When would the make or buy decision involve investment analysis? How would you proceed in this case?

Q4-7. In the materials handling case, why was it necessary to employ simulation to determine the benefit/cost ratio?

Q4-8. What macro design problems are likely to be encountered in the following situations?

- (*a*) Court system
- (*b*) High school curriculum
- (*c*) Commercial bank
- (*d*) Hospital pharmacy

Q4-9. What noneconomic factors might influence the choice of a transformation process for the manufacture of a new electronic kitchen oven?

Q4-10. Does the routing process lend itself to becoming a part of a computer-based information system? If so, what would be the major characteristics of such a system, and how could it be used?

Q4-11. If you were employed by the consulting firm that was commissioned to study the Smithdale Hospital problem, how would you have designed the simulation experiment?

Q4-12. A manufacturer has decided to add a line of outboard motors. What factors might he consider in making a choice between a continuous- and an intermittent-production process?

Q4-13. Suppose that the marginal manufacturing costs for toroids in the Sanbourne Company case was $4. What would your decision be—make or buy?

Q4-14. Suppose that a fixed conveyor system is to be installed whose purpose it is to move work-in-progress from one work station to the next. The work will be attached to the conveyor so that the necessary operations can be performed as the job is moving. What is the consequence of a slower conveyor speed together with less spacing between jobs? How could you test different speed and spacing strategies?

Q4-15. Is work flow design related at all to human needs? What role should these needs play in the macro design of the system?

Q4-16. In the past the U.S. Post Office has maintained labor-intensive methods for the processing of mail. How might the system be automated? Would you foresee any implementation problems?

PROBLEMS

P4-1. A casting for a new product must be processed on one of three machines. The first alternative is a numerically controlled machine with a fixed setup cost of $225 and a per-piece cost of 20 cents. The second alternative is a press with a setup cost of $150 but a per-piece cost of $1. The third is a lathe with a setup cost of $60 and a per-piece cost of $2. If the anticipated volume per order is 150 units, which alternative should be chosen?

P4-2. Suppose that the manufacturing manager who must make the decision depicted in Prob. P4-1 is uncertain as to the actual level of future orders. He feels that there is a 10 percent chance that the orders will be for 50 units, a 60 percent chance that the orders will be for 100 units, and a 30 percent chance that they will be as large as 150 units. Suggest several criteria which he may use to solve his problem; solve, and compare the results for each one.

P4-3. Draw a product flow process chart for the following process:

(*a*) Raw stock for subassembly A delivered to cutting table
(*b*) Wait for next
(*c*) Stock cut to size
(*d*) Sent to press
(*e*) Wait for next
(*f*) Press setup
(*g*) Press operations
(*h*) Inspection
(*i*) Send to plating shop
(*j*) Chemically cleaned
(*k*) Electroplated
(*l*) Inspected
(*m*) Raw stock for subassembly B delivered to cutoff saw
(*n*) Wait for next
(*o*) Stock cut
(*p*) Send to drill press
(*q*) Wait for next
(*r*) Stock drilled
(*s*) Inspected
(*t*) Assembly of subassembly A to B
(*u*) Inspect
(*v*) Send to stockroom

P4-4. Suppose that the second alternative which the Grandiron Company is considering is the purchase of a much faster forklift to replace its present one. This faster truck will cost $19,000 and have a life expectancy of 10 years. Its estimated service times are given below.

Time	*Likelihood*
3	.15
4	.25
5	.30
6	.30

Perform one replication and compute its benefit/cost ratio.

Note: It would be more effective if each student started from a different place in the random number table and upon completion of these replications a benefit/cost distribution could be constructed on the board.

P4-5. Determine a benefit/cost ratio for the material handling case presented in the chapter, given the following arrival and service patterns. Perform only one replication and assume that the other parameters are unchanged.

Time between arrivals

Time	*Number of occurrences*
4	6
5	19
6	15
7	35
8	25
	100

Service time

Time	*Number of occurrences*
4	15
5	35
6	30
7	20
	100

BIBLIOGRAPHY

Buffa, Elwood S.: *Modern Production Management,* John Wiley & Sons, Inc., New York, 1969.

Ganett, J. William: *Production and Operations Management,* Harcourt Brace Jovanovich, Inc., New York, 1968.

Timms, Howard L., and Michael F. Pohlen: *The Production Function in Business,* Richard D. Irwin, Inc., Homewood, Ill., 1970.

5

Process Design: Micro

INTRODUCTION

In the previous chapter the focus was on the macroscopic problem of process design. At that stage we examined whether the process should be labor- or capital-intensive, the kinds of equipment that would be needed, the flow of materials through the process, the make or buy decision, and the investment in a materials handling system. It should be reiterated that these conclusions are not necessarily final. For, as the design process continues and the focus shifts to the micro design stage, some redesign in the macro stage might prove necessary.

The micro design problem might also be called the *job design* or *work station design* problem. The work station becomes the center of attention, and analysis focuses on the steps that should be included in a particular station as well as on the methods for executing these steps.

In totally automated and highly capital-intensive industries such as chemicals and petroleum, the process design is rigidly specified by manufacturing technology and the individual plays only a small role in the production process. His only interaction is as a control agent. Therefore in the micro design phase of these industries, the concern is limited to such issues as the proper position of equipment controls and the ability to read them quickly and accurately and to respond in a short period of time. In addition the sequence of steps which the operator must perform during each cycle should be carefully designed to ensure a smooth flow.

Where automation is not so complete, the individual plays an integral role in the *production* and *control* functions associated with the set of steps at his work station. At his station we find a set of steps that have been assigned to make up the job. In some cases the job is strictly labor-intensive and we find no tools at the work station. In others the operator may be utilizing one or several tools.

In the next section we will consider an overview of the problems which arise when individuals play a significant role in the production process. Then we will look at three of these issues in greater detail. They include the balance of assembly lines, the establishment of effective work methods, and the setting of labor standards. Often it is these three issues that dominate the micro design problem.

AN OVERVIEW OF THE MICRO DESIGN PROCESS

PRECEDENCE RELATIONS

Quite often the first problem encountered in the micro design stage is the determination of the particular set of steps that should make up the job content of the work station. Usually the different steps which comprise the entire task can be executed in one of several different sequences. Consider the job design problem which faces the Getwell Electronics Company. Their new blood analyzer requires several identical printed-circuit boards. Based on a recent forecast for the demand of these analyzers, approximately 170 printed-circuit boards should be assembled and soldered per day.

Getwell's industrial engineers have identified the steps necessary in the assembly process, their precedence relationships, and estimates of the standard time that each of these steps should take. All of this information is found in Fig. 5-1.

Perhaps a few words ought to be said about the origin of this diagram. The precedence relationships were identified during the macro process design phase. At that stage the printed-circuit board with which we are now concerned was but one small part of the complete production job. Once the focus shifted to the micro aspects, each subassembly was considered separately.

STANDARD TIMES

The standard times which appear in Fig. 5-1 were obtained from their company standards and represent the time in which these motions can be completed by *most* people.

Emphasis was placed on *most* because variability is inherent in the human process. For example, the average time to mount a transistor might be 25

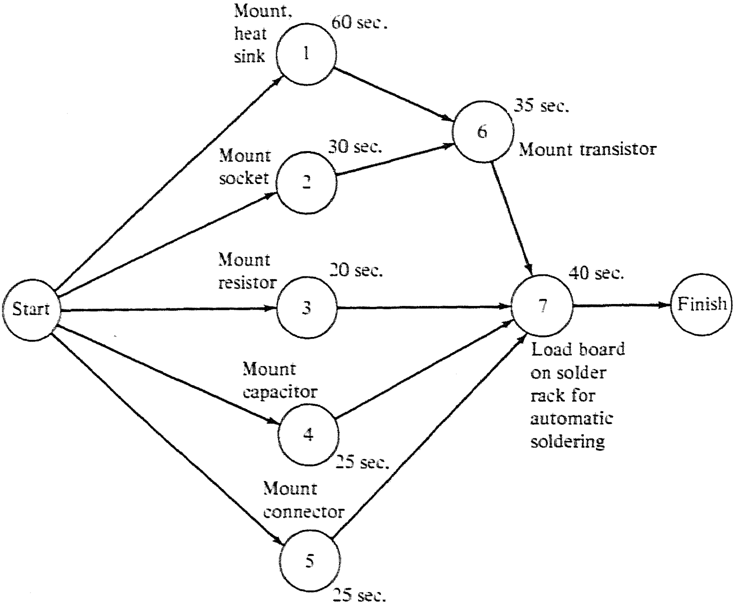


Fig. 5-1 Precedence chart for printed-circuit board.

seconds; but if this is used as a standard time, approximately half of the people performing the task would take longer. Of course, half would take a shorter time. Perhaps a figure such as 35 seconds should be used. Most should be able to meet this standard.

LINE BALANCING

Once the standard times are determined, the next action is to group these steps into jobs. Since 170 boards per day are required and each board takes 235 seconds to complete, approximately 39,950 seconds of production time would be required. With two operators approximately 2 × 3,600 seconds/hour × 6 hours/day = 43,200 seconds per day are available. Therefore we can assume that two full-time operators will be utilized to accomplish this objective.

How will these steps be assigned to the two work stations? This depends on the criteria by which the Getwell Electronics Company designs a process. If that criterion is simply the minimization of costs, perhaps the steps should be split between these two work stations to take advantage of the inherent efficiencies associated with job specialization. This might be accomplished by assigning one set of steps to station A and another set to station B. One alternative assignment might be as shown in Table 5-1.

This particular division of steps does not result in a perfectly balanced set of work stations. Station B will be idle during some phase of each cycle. Perhaps another division of steps might result in a better balance. The ideal,

Table 5-1 Assignment of steps to stations A and B

Station A		*Station B*	
Step	*Time*	*Step*	*Time*
1	60	3	20
2	30	4	25
6	35	5	25
	125 seconds	7	40
			110 seconds

of course, would be for each work station to have exactly half of the work with no idle time.

OPERATOR VARIABILITY

Even if the line were perfectly balanced, there would have to be an in-process inventory between stages. Remember that the production times which are given represent standards which can be met most of the time. However, because of the variability of operator performance, there will be occasions when station A will *exceed* the standard production time and, without in-process inventory, station B would be out of work. It could also happen that station B's production time will be much *less* than the standard, and on this occasion production will also cease without in-process inventories. The appropriate size of in-process inventories is a difficult problem to solve.

Let us take a closer look at this problem. Figure 5-2 illustrates a three-work-station, semipaced production line. A *semipaced* line is a line in which work stations are connected in tandem and in which the input of one station depends upon the output of the previous one. The work pieces, however, are *not* fixed to a moving line. In a *paced line,* the work pieces are fixed to a moving line. The printed-circuit assembly represents a semipaced problem.

Now if each of these stations were perfectly balanced at 2 minutes each, one would not expect to find a finished part emerging from station 3 every 2 minutes. The rate at which parts would be generated would depend upon the within-station variability. If a part is to be finished on the average of every 2 minutes, an in-process inventory must be maintained between these work stations.

One way to determine the adequacy of in-process inventories is to simulate the production process for several different levels of inventory. To begin with,

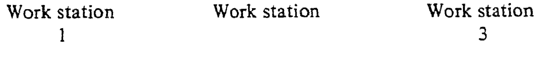


Fig. 5-2 Semipaced production line.

a probability distribution which describes station performance is required for each station in the sequence. This can be obtained in one of three ways. The first would be to use actual production data and compile a discrete probability distribution for each station. The second would be to develop a subjective probability distribution based on educated guesses, and the third would involve using standard mathematical distributions such as the Poisson. The next step would be to choose an inventory strategy for each station. Then the simulation could proceed. Work would be processed through the system and the rate of output would be observed. This process can then be repeated for different levels of inventory. What is usually seen is that the rate of output rapidly increases as the inventory is increased from zero. These inventory increases, however, seem to reach the point where they have less and less of an effect on output.

In-process inventories cost money to maintain. The money tied up represents funds which are therefore not available elsewhere for the purpose of contributing to the profitability of the firm. Larger in-process inventories, therefore, may make a small contribution to the output rate of the line, but they may represent an inefficient use of funds.

If in the macro design phase Getwell Electronics Company had decided upon a paced line, they would have had to consider other means of solving the variability problem. Under that strategy the length of time that the operator has in which to perform his work is constant from one work piece to the next. Since the operator's pace varies from one piece to the next, job cycle time, or the length of time that the work piece is available, must be set well above the average operator time. Otherwise some work pieces would proceed down the paced line unfinished or finished hastily and poorly. Figure 5-3 shows the relationship between operator time and cycle time. As cycle time is increased, parts will go unfinished a smaller percentage of the time, but the output rate of the line also decreases.

There are a few remedies for this problem of variability. One is to increase the length of the work station so that the operator always has access to more than one work piece. In this way, if one work piece takes longer than average, the operator has a chance to make up the extra time on the next piece. Another remedy is to decrease the space between the work pieces and at the same time decrease the speed of the conveyor. Again, whatever time is lost on one piece can possibly be made up on the next.

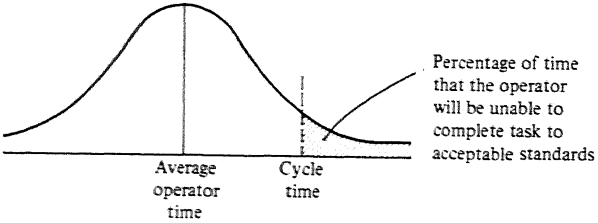

Percentage of time that the operator will be unable to complete task to acceptable standards
Average operator time
Cycle time

Fig. 5-3 Relationship between operator time and cycle time.

WORK METHODS

Once the set of steps to be included at each work station has been determined, Getwell Electronics Company must shift its focus to the selection of a method for executing these tasks. The use of assembly aids and fixtures, the physical design of the work station, and the assembly techniques used by the operator must all be considered.

Where volume warrants, a significant amount of experimentation can be undertaken to ensure that the most efficient strategy is being employed. During this phase it is not uncommon to see the use of "left- and right-hand" charts to record the operator's motions, the use of high-speed motion picture films, and even the use of simulation techniques. They act as vehicles for carefully analyzing each work method alternative. Since the volume anticipated by Getwell Electronics Company is not large, perhaps only the "left- and right-hand" chart could be used to help establish work methods.

LABOR STANDARDS

After the line has been balanced and the work methods have been finalized, all of the details for performing each job are recorded as labor standards. These standards can then be used for performance evaluation purposes, training of new operators, cost control, and manpower planning.

JOB SPECIALIZATION VERSUS JOB ENLARGEMENT

Frequently the criteria for job design include not only economic but human factors. These human factors center around the individual's role in the transformation process. Perhaps the most popular effort in this direction is the job enlargement school. The belief of this school is that the individual plays a crucial role in the effectiveness of the production process, and that, therefore, not only physiological but psychological and sociological characteristics must be considered in the design process. Consequently their focus is on the job itself, and their hypothesis is that if the job is satisfying, the individual will be properly motivated.

If in the printed-circuit example we were interested in enlarging the job, we might assign all seven steps to each work station. Perhaps we might even make the operators totally responsible for the finished product by requiring them to inspect the unit and stamp it with their employee number.

Just how beneficial job enlargement can be is an issue with pros and cons. In the next chapter we look more closely at this problem and in general at the role that human factors should play in the design of jobs.

LINE BALANCING

Line balancing is a rather complex problem that crops up in both paced and unpaced processes. Much work of a theoretical and practical nature has been directed at this problem. In the next section an integer programming model is presented which in the final analysis proves to solve the problem perfectly but which is too time-consuming, even on the fastest of computers. The reason for presenting it here, however, is to give you an idea of one theoretical approach to the full-blown problem. Those not wishing this exposure can move to the following section, which focuses on a more practical approach to solving the problem. In that section a good, but not the best, solution to the problem is presented. Before jumping into these solution procedures, let us look at the general form and magnitude of the line-balancing problem.

The essence of the line-balancing problem can be seen in Fig. 5-1. Here we find a number of steps, some of which can be performed only after others have been completed. The objective is to distribute the steps evenly over a set of work stations so that idle time is minimized.

The number of ways in which these steps can be placed into a feasible sequence is quite large. In general, N steps can be placed in $N!$ sequences. However, because of the precedence relations, only some of the $N!$ sequences are feasible. For r precedence relationships there are approximately $N!/2^r$ feasible sequences. Therefore in Fig. 5-1 there are approximately $7!/2^6 = 78$ feasible sequences. As the number of steps increases, the number of feasible sequences goes quickly out of sight. For example, with 70 steps and 105 precedent relations, the approximate number of feasible solutions is $70!/2^{105} = 10^{65}$. How many centuries would it take all the computers to enumerate these possible sequences? Many. Therefore, rather than complete enumeration, we must turn to more efficient means of establishing a solution.

//// LINEAR PROGRAMMING MODEL

In general, most models which have been developed for solving the line-balancing problem require as data a cycle time C. We define *cycle time* as the length of time that the work piece is available at each work station. Therefore, before the analysis can proceed, it must be determined whether a short cycle time or a long cycle time is preferred. This of course depends upon the firm's attitude toward job enlargement versus job specialization. Once determined, the line-balancing effort can proceed.

If the total time required to complete a job, temporarily ignoring precedence constraints, is T, the minimum number of stations is $S_{\min} = T/C$. However, since precedence constraints must be observed, it might be impossible to perform the job with so few stations. At worst the maximum number required, $S_{\max}$, will be equal to the number of steps in the job. Frequently, however, it will be obvious that the maximum number will be less than this. Although there is no problem in specifying a maximum number much greater than is necessary, there is a problem if this number is not large enough. What occurs is that it will be impossible to assign the steps to those stations without violating the cycle time constraint. Our objective in solving this problem is to distribute the steps so as to minimize the number of stations required, S_{bal}, where

$$S_{\min} \leq S_{\text{bal}} \leq S_{\max}$$

and at the same time not exceed the specified cycle time C or the precedent constraints.

Consider the problem shown in Fig. 5-4. Suppose that the production objective for this job is three finished pieces per hour or one every 20 minutes. One strategy might be to assign the entire 75-minute job to one person. This being the case, each worker would produce 5.6 units during a 7-hour working day. To meet the daily quota of 21 pieces, four stations would be required.

Another strategy would be to assign 20 minutes' worth of work to each station. The layout of the facility would be such that the output from one station would become the input to the next. At least four stations would be required, $S_{\min} = 4$, whereas the maximum number of stations $S_{\max}$ would be 8. Now we will proceed to determine the exact number of stations required as well as the particular station to which these steps must be assigned.

$$\begin{aligned} \text{Let } i &= \text{the } i\text{th step} \qquad & i &= 1, 2, 3, \ldots, N \\ j &= \text{the } j\text{th station} \qquad & j &= 1, 2, 3, \ldots, S_{\max} \end{aligned}$$

Then t_i = time required to complete ith step

$$X_{ij} = \begin{cases} 1 & \text{if step } i \text{ is assigned to station } j \\ 0 & \text{if step } i \text{ is assigned to some other station} \end{cases}$$

Cycle time constraints First we formulate a set of constraints which limit the number of steps assigned to each station so that the cycle time is not violated.

$$\sum_{i=1}^{N} t_i X_{ij} \leq C \qquad \text{for } j = 1, 2, 3, \ldots, S_{\max}$$

Applying this to our problem, we have

$$\begin{aligned} &11X_{11} + 17X_{21} + 9X_{31} + 5X_{41} + 8X_{51} + 12X_{61} + 3X_{71} + 10X_{81} \leq 20 \\ &11X_{12} + 17X_{22} + 9X_{32} + \cdots + \cdots + \cdots + \cdots + 10X_{82} \leq 20 \\ &\cdots \\ &11X_{18} + 17X_{28} + 9X_{38} + \cdots + \cdots + \cdots + \cdots + 10X_{88} \leq 20 \end{aligned}$$

where the first constraint limits the number of steps which can be assigned to station 1 so that the time required to perform these steps does not exceed

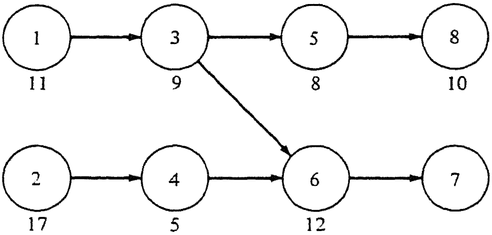


Fig. 5-4 Precedence chart for line-balancing problem.

the cycle time of 20 minutes. The second constraint limits the steps assigned to station 2, and so on.

Step constraints We must also specify a set of constraints that makes sure that every step is assigned to one and only one station.

$$\sum_{j=1}^{S_{\max}} X_{ij} = 1 \qquad \text{for } i = 1, 2, 3, \ldots, N$$

For this problem we have

$$X_{11} + X_{12} + X_{13} + X_{14} + X_{15} + X_{16} + X_{17} + X_{18} = 1$$

$$X_{21} + X_{22} + X_{23} + X_{24} + X_{25} + \cdots + X_{27} + X_{28} = 1$$

$$X_{81} + X_{82} + X_{83} + \cdots + \cdots + \cdots + X_{87} + X_{88} = 1$$

where the first constraint ensures that step 1 is assigned to only one of the possible eight stations. The second constraint ensures that step 2 is assigned to only one of the possible eight stations, and so on.

Precedence constraints Next we must express the precedence relationships as linear constraints. To accomplish this we have only to consider a step's relationship to its preceding step. Consider the precedent relationship shown in Fig. 5-4. Step 3 must follow step 1, and this can be written as

$$X_{31} \leq X_{11}$$

Therefore, if step 1 is assigned to station 1, then $X_{11} = 1$; and if step 3 is not assigned to station 1, then $X_{31} = 0$. Since $0 \leq 1$ is feasible, this distribution is allowed. What if step 1 were assigned to station 2 and step 3 were assigned to station 1? Then $X_{11} = 0$ and $X_{31} = 1$ and $1 \nleq 0$. Therefore it would not be feasible.

Continuing with the remainder of the precedent relationships, we have

$$X_{32} \leq X_{11} + X_{12}$$

$$X_{33} \leq X_{11} + X_{12} + X_{13}$$

$$X_{78} \leq X_{61} + X_{62} + X_{63} + X_{64} + X_{65} + X_{66} + X_{67} + X_{68}$$

$$X_{88} \leq X_{51} + X_{52} + X_{53} + X_{54} + X_{55} + X_{56} + X_{57} + X_{58}$$

Objective function The purpose of the objective function will be to push the final steps, 8 and 7, into previous work stations. Why do we have to "push" only 8 and 7 into earlier stations? Because if they are successfully pushed, all the other steps are pushed through their precedence constraints.

We established at the beginning of our analysis that at least four stations were required. We therefore need to focus only on stations 5, 6, 7, and 8. To push the last steps into earlier stations, we make an assignment of these steps to a later station much more costly than their assignment to an earlier one. We can accomplish this by specifying the objective function in the following way:

$$\text{Minimize } Z = 1(X_{75} + X_{85}) + 10(X_{76} + X_{86}) + 100(X_{77} + X_{87}) + 1{,}000(X_{78} + X_{88})$$

Essentially we have generated an arbitrary cost explosion from station 5 to station 8.

The X_{ij}'s are decision variables and as a result should be restricted to a value of 0 or 1. Fractional values would imply that the step was to be split between two work stations. This, of course, makes little sense. Therefore the problem should be considered as integer and the appropriate integer programming computer code used for its solution.

What has probably become obvious to you is that a considerable number of equations are necessary to solve this eight-step problem for the exact solution. In a more realistic problem with hundreds of steps, the integer programming formulation would get out of hand. In almost all cases this burdensome effort to get the exact solution is not worth the extra effort when compared with methods that get good although not necessarily the best answers. We now turn to one such method. ////

A HEURISTIC APPROACH TO LINE BALANCING

The line-balancing model to be presented here is considered a heuristic rule.[1] A *heuristic* is a rule of thumb which may have little, if any, theoretical foundation but is found to generate very good although not necessarily optimal solutions.

The problem that will be solved by this heuristic can be found in Fig. 5-5. The first step in the solution process is to redraw the precedence diagram. First, columns are established under which the steps will be entered. Six columns are specified in Fig. 5-6. Next the steps are drawn into the columns as far to the left of the diagram as possible. Care must be taken not to violate the precedent constraints. Figure 5-7 shows a common error in setting up the new diagram.

Two basic ideas are used in the assignment of steps to stations. These include the permutability within a column and the lateral transferability of steps across the chart. The first concept, that of permutability within a column, is the freedom to select any combination of steps in a particular column. In

---

[1] See L. Wester and M. D. Kilbridge, "Heuristic Line Balancing: A Case," *Journal of Industrial Engineering,* vol. 13, no. 3, May–June 1962.

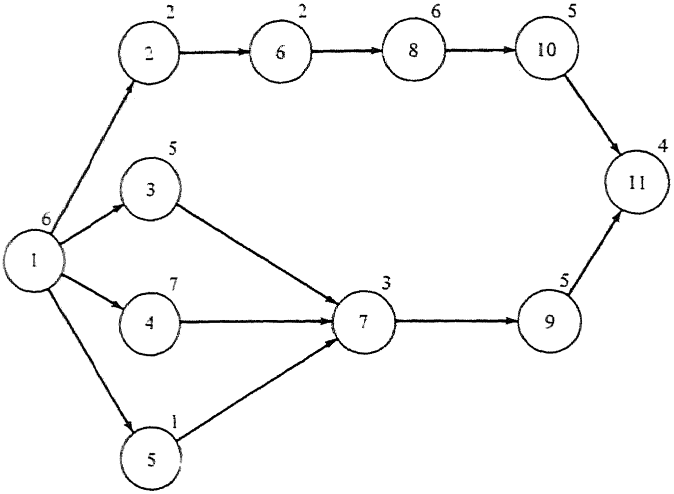


Fig. 5-5 Line-balancing example.

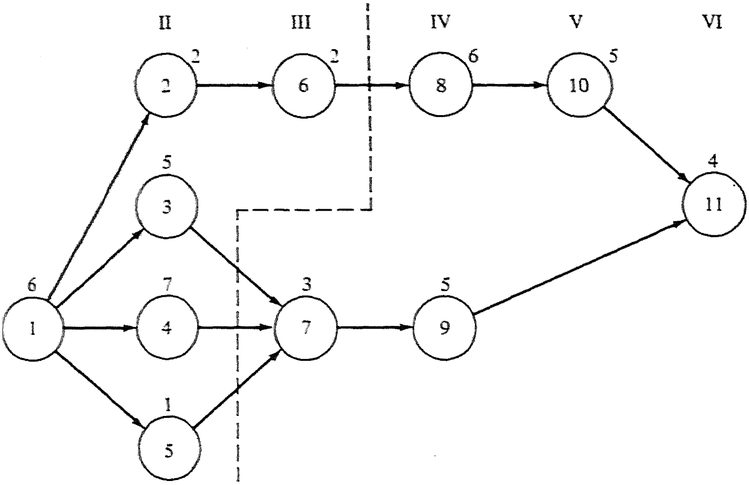
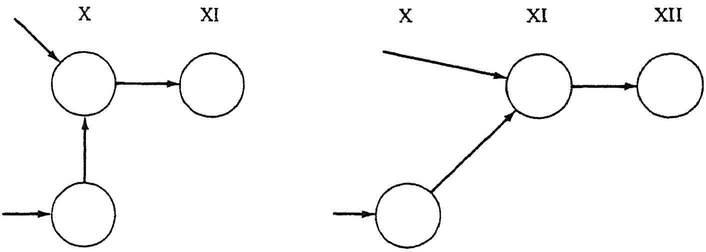


Fig. 5-6 Steps drawn below columns and moved as far to the left as possible.

our example we will start with the objective of balancing a line given a cycle time of $C = 23$ minutes. First we proceed down the first column. Only one step appears in this column, and it has a time of 6 minutes. We include this step in the first station and proceed to column II. It is important to note that only after all the steps in a particular column have been assigned can you move to the next column.

Moving to column II, we find that if all of the steps are placed in station



Not acceptable

Acceptable

Fig. 5-7 A common pitfall.

1, the total time thus far assigned is 21 minutes. Now we must move to column III to complete the 23-second cycle time.

Since 21 minutes' worth of work has already been assigned, we will try to select some combination of steps from column III to equal the $23 - 21 = 2$ minutes necessary to balance station 1. Of course, it might not be possible to find a combination that would exactly equal 2 minutes, in which case the combination which came closest would be selected. It is here that use is made of the first concept, permutability within a column, for any combination of steps from a column which does not exceed the desired time is allowed.

In this case the choice is easy. We find a step with a time of 2 minutes, and the first station can be perfectly balanced. A dotted line is used to set off the included steps in station 1.

Turning to station 2, the remaining 3 minutes associated with step 7 are assigned to station 2. The entire contents of column IV can be placed in station 2, as well as the contents of columns V and VI. In fact, station 2 is also perfectly balanced with a total time of 23 minutes. Both stations have been perfectly balanced.

ANOTHER EXAMPLE

Now let us try to balance the line if the cycle time is set to $C = 10$ minutes. The precedent chart is redrawn in Fig. 5-8. First, step 1 is assigned to station 1. There are now $10 - 6 = 4$ minutes to be assigned to station 1. If steps 2 and 6 could be assigned to station 1, then it would be perfectly balanced. However, according to our rules, step 6 can be assigned only after every step from column II has been assigned. Now the second concept of lateral transferability can be illustrated. Steps can be moved laterally as long as the precedent relationships are not violated. For example, steps 3, 4, and 5 can be moved into column III provided that steps 7 and 9 can be moved one column to the right. It can be seen from Fig. 5-8 that this is indeed possible. The move is made in Fig. 5-9. Now step 2 is included in station 1 and selection

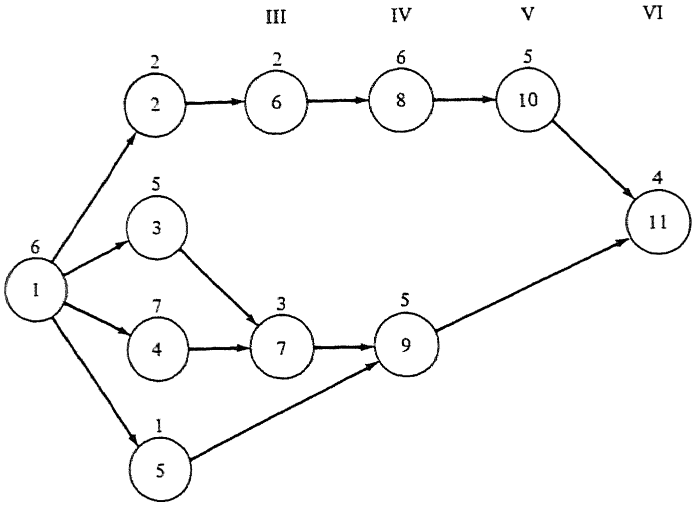


Fig. 5-8 Line-balancing example.

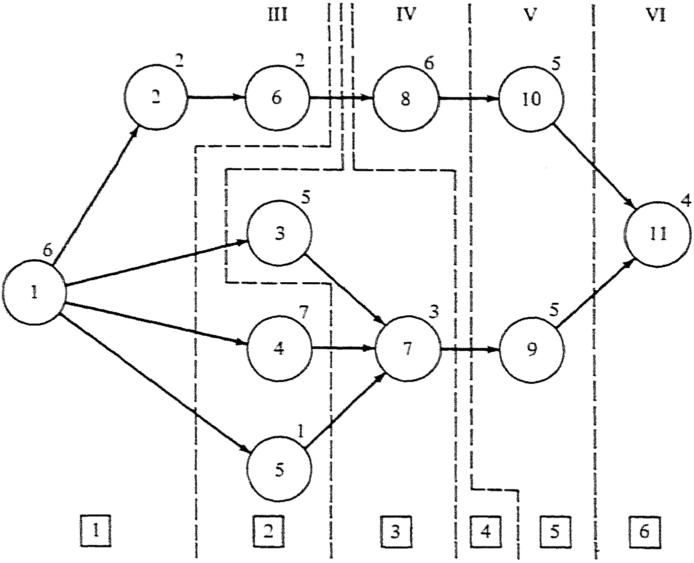


Fig. 5-9 Solution.

can proceed to column III, from which step 6 is included to perfectly balance station 1 at 10 minutes.

Next we proceed to assign steps to station 2. Since lateral transferability is no longer possible, steps 4 and 5 are assigned to station 2. Steps 3 and 7 can be assigned to station 3, step 8 to station 4, steps 9 and 10 to station 5, and step 11 to station 6.

This heuristic, which was developed by Kilbridge and Wester, has proved to be a simple and powerful technique across a wide spectrum of line-balancing problems. There are some situations, however, where short cycle times are accompanied by *the presence of two or more stations* assigned to one column when this technique does not perform as well as some others. So once again we see that one model is not general enough to be used indiscriminately.

WORK METHODS

Once the optimum combination of steps for each work station has been determined, focus must shift to the optimum way in which this set of steps should be carried out. The effort directed at designing efficient work methods depends, of course, on the production quantity. As production quantity increases, the design of work methods becomes more sophisticated and the techniques of analysis progress from the informal to the use of high-speed motion picture cameras.

OPERATION CHARTS

In the simple case where the cycle for the job is short and the production volume is low to moderate, *operation charts* are used. The purpose of this type of chart is to show the coordination of movements between the left and right hands during an assembly sequence. As you can see from Fig. 5-10, there are two columns, one for left-hand and another for right-hand movements. The motions are broken down into reach, position, grasp, transport, and assemble. Breaking these general motions down into finer elements for jobs of low to moderate volume is usually considered unnecessary.

In Fig. 5-10 the motions are shown for an initial design sequence of a bolt and washer assembly. The analyst, after developing a design sequence, should carefully scrutinize it for motions which might be fatiguing or wasted. Perhaps a redesign based on the use of new tools or fixtures could result in a better balance between the hands; frequently a better balance means a shorter cycle time. A useful guide when performing this analysis is Frank B. Gilbreth's[2] set of principles of motion economy. It provides valuable insight into the efficient use of the hands, arrangement of the work place, and design of tools and equipment.

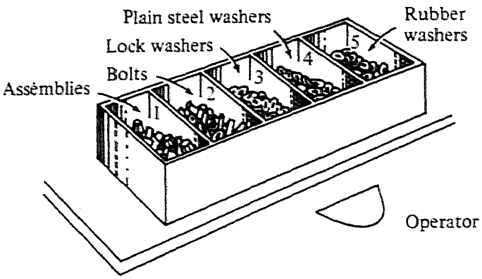
Putting some of these principles to use, a redesign of the bolt and washer assembly was undertaken. If a fixture is designed to hold two assemblies, then both can be assembled at the same time. Figure 5-11 illustrates the new work method, as well as the fixture, which is mounted in front of the operator.

MICROMOTION ANALYSIS

If the volume warrants, a micromotion analysis is undertaken, where each job is broken down into very basic elements called *therbligs*. This is a much finer breakdown than in the operation chart. The activities of the left and right hands are plotted against a time scale, and the data for this "simo chart" are

---

[2] Frank B. Gilbreth, *Primer of Scientific Management*, D. Van Nostrand Company, Inc., New York, 1914.



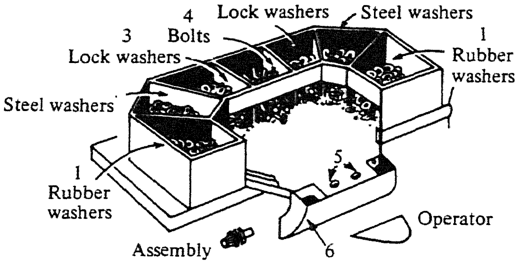
Left Hand		Right Hand
Carries finished assembly to bin 1		Reaches for lock washer in bin 3
Releases assembly into bin 1		Grasps lock washer from bin 3
Reaches for bolt in bin 2		Carries lock washer to central position
Grasps bolt from bin 2		
Carries bolt to central position		Positions lock washer
		Assembles lock washer onto bolt
		Reaches for plain steel washer in bin 4
		Grasps steel washer from bin 4
		Carries steel washer to bolt
		Positions steel washer
Holds bolt	○	Assembles steel washer
		Reaches for rubber washer in bin 5
		Grasps rubber washer from bin 5
		Carries rubber washer to bolt
		Positions rubber washer
		Assembles rubber washer
Carries finished assembly to bin 1	○	Releases finished assembly

Fig. 5-10 Operation chart of bolt and washer assembly. (*From Ralph M. Barnes, Motion and Time Study, John Wiley & Sons, Inc., New York, 1964, p. 113.*)

usually obtained with a high-speed camera. The pictures are taken at approximately 1,000 frames per second, and when they are replayed at 60 frames per second, every detail in the process can be studied carefully.

MAN/MACHINE CHART

When the operator at a work station must interact with one or several machines on a periodic or random basis, an activity analysis is performed. Again the



Left hand	Right hand
Reaches for rubber washer in bin 1	Reaches for rubber washer in bin 1
Grasps rubber washer from bin 1	Grasps rubber washer from bin 1
Slides rubber washer to countersunk hole	Slides rubber washer to countersunk hole
Positions rubber washer in countersunk hole 5	Positions rubber washer in countersunk hole 5
Reaches for plain steel washer in bin 2	Reaches for plain steel washer in bin 2
Grasps steel washer from bin 2	Grasps steel washer from bin 2
Slides steel washer to countersunk hole	Slides steel washer to countersunk hole
Positions steel washer in countersunk hole 5	Positions steel washer in countersunk hole 5
Reaches for lock washer in bin 3	Reaches for lock washer in bin 3
Grasps lock washer from bin 3	Grasps lock washer from bin 3
Slides lock washer to countersunk hole	Slides lock washer to countersunk hole
Position lock washer in countersunk hole 5	Positions lock washer in countersunk hole 5
Reaches for bolt in bin 4	Reaches for bolt in bin 4
Grasps bolt from bin 4	Grasps bolt from bin 4
Carries bolt to washers at 5	Carries bolt to washers at 5
Positions bolt preparatory to inserting it into washers at 5	Positions bolt preparatory to inserting it into washers at 5
Assembles bolt and washers	Assembles bolt and washers
Lifts bolt and washers, carries to left and releases into top of chute 6	Lifts bolt and washers, carries to right and releases into top of chute 6

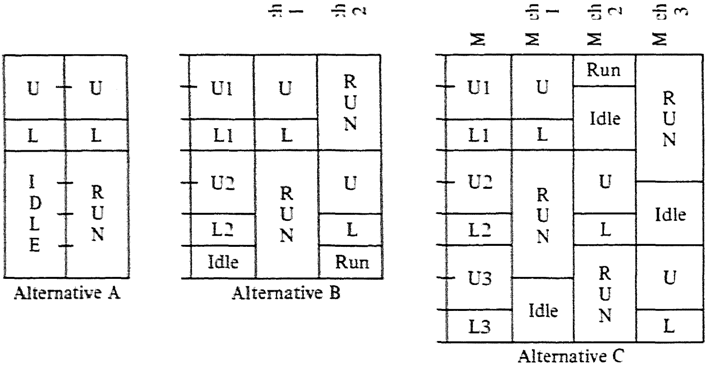
Fig. 5-11 Operation chart of bolt and washer assembly—redesign. (*From Ralph M. Barnes, Motion and Time Study, John Wiley & Sons, Inc., New York, 1964, p. 114.*)

central coordinating axis is time, but in this case the individual's activity is recorded on one side and the state of the machine on the other. The object of this analysis is to scrutinize the activity of the man and the machine in an attempt to achieve the most economical balance of activity. For example, if excessive operator idle time became apparent as a result of this analysis, perhaps he might operate two identical machines making the same product or he might even operate another machine making a different product. Even if the resultant job redesign leads to some machine idle time, it could still be economically superior to the old method. However, before this can be proved conclusively, an economic analysis of the before-and-after situation must be undertaken.

In Fig. 5-12 an analysis is shown for each of three alternatives. These include an operator running one, two, and three machines. It can be seen that the per-piece cost for alternative B is the lowest; therefore the most economical solution is to have the operator run two machines.

This simple deterministic analysis is quite appropriate when the time required for each of these activities is known with certainty, or when the certainty assumption is a good approximation.

There are occasions, however, when the time for carrying out these activities is subject to considerable uncertainty. Then a stochastic model must be built. To illustrate this, we turn to a problem found in hospital administration.



U = unload
L = load

Machine cost per hour = $5
Labor cost per hour = $10

Machine output rate = 1 piece per cycle

Cycle time	7 min	7 min	9 min
Man idle	57%	14%	0%
Machine idle per machine	0%	0%	22.2%
Pieces per hour	$\frac{60}{7}$	$\frac{120}{7}$	20
Per-piece cost	$\frac{15}{60/7} = 1.75$	$\frac{20}{120/7} = \frac{7}{6} = 1.166$	$\frac{25}{20} = 1.25$

Fig. 5-12 Three alternative man/machine assignments.

STOCHASTIC MAN/MACHINE MODELS

A stochastic problem can be found in the field of hospital administration whose form is quite similar to the man/machine problem but whose content is quite different. The problem is to determine the number of patients who can be assigned to one nurse. Certain constraints must not be violated. For example, if we break up all calls for service into critical and noncritical, a noncritical call must wait no longer than, say, 7 minutes, while a critical call must wait no longer than 40 seconds. One technique which might be used to evaluate a set of alternative solutions is simulation. Several different combinations of nurses and patients could be simulated, one at a time. In a simple model only time between calls, percent of critical calls, and service time need be considered in the analysis. The results of each simulation would show the average waiting time for critical and noncritical calls and the number of times that the constraints were violated. The best proportion of patients to nurse would be the one which did not violate the constraints at all or the one which had the fewest violations.

This analysis could be extended to determine nursing loads and assignments on a continuing basis. First the patient load together with expected nursing-staff levels would have to be kept up to date in the computer's file. Then as patients are admitted to the hospital, their health status, room assignment, and expected level of nursing care would be entered into the computer. A simulation could be run before every nursing shift, determining assignments, in such a way as to minimize the number of service violations.

PHYSIOLOGICAL FACTORS

Another aspect of the man/machine system is the way in which the machines and equipment interface with the human operator. Until the Second World War, little attention was focused on the role that the operator played in carrying out his activities. Then with the advent of complex military equipment it became important to consider the effective working space of the operator, the efficient design of hand and foot controls, the location of displays, and the effective use of the hands and body. Newer aircraft, for example, were designed in such a way that cockpit controls were positioned within comfortable reach and with the most frequently used controls the easiest to reach. The shapes of hand controls were differentiated so that pilots could operate by feel alone, without having to visually check to ensure that the proper control was being operated. Instruments were positioned in such a way that in the normal operating position all controls read in the same direction. In this way, if the instruments read "12 o'clock," all systems were normal. A quick glance could therefore determine the state of the instruments.

The outgrowth of this World War II effort was the application of these ideas to the industrial sector of the economy. These data are now used to help in the layout of the work station and in the design of production machinery. They serve as guides in designing the maximum and normal working areas and the heights for chairs and tables. They also represent well-established standards for the strength and forces of different body movements, the speed and precision associated with movement of body parts, and the most effective way to design visual displays.

THE WORK STATION ENVIRONMENT

It is interesting to see how the focus has shifted from one set of environmental problems to another in the last 40 years. One of the original goals of the Hawthorne study was to determine the effect of lighting on worker productivity. Lighting and heating were seen as having short-run effects; if the levels went below some threshold, activity fell off sharply. For the most part, problems due to improper lighting, heating, and humidity are virtually nonexistent in our modern society. However, as these short-run impediments to worker productivity have been overcome, the emphasis has shifted to the longer-run effects of noise and atmospheric pollution found *inside* many plants. Both of these sources of "pollution" do not result in harmful short-term effects, and as a result business has had little economic incentive to alleviate these conditions. Management's future course of action, however, must be based on moral rather than economic principles; and if this responsibility is not shouldered, we should not be surprised to find the enactment of stricter state and federal regulation. Already the Walsh-Healy act of 1969 requires that those entering into contracts exceeding $10,000 with any agency of the United States comply with federal noise standards.

LABOR STANDARDS

Labor standards play an important role in the design and operation of any system. In some industries the standards are developed and applied on a very informal basis; in others, considerable effort is undertaken to develop standards, make use of them, and readjust them periodically. An example of the informal process can be seen in small production shops, retail stores, and service industries. The formal process is seen in large firms where the manufacturing volume is high and many employees are involved in repetitive activities.

The uses of labor standards are many. They are used in the development of wage incentive plans for the costing and pricing of products, the planning of manpower requirements, the process design phase as the criterion for comparing alternative processes, and, as we have already seen, as input data to the line-balancing problem. They are also used in the cost control process for comparison with actual costs. For example, if the actual cost deviates substantially from the standard cost, corrective action may be taken.

Three methods for establishing work standards will be presented. These include the time study approach, the predetermined time study approach, and work sampling. The first requires that we time the actual job, using a stopwatch, whereas the second makes use of a set of predetermined times that have been well tested for a host of standard motions. The third approach is used for jobs which have a fairly long cycle time. Typically it is used for setting standards for clerks, secretaries, bank employees, hospital personnel, government

Time study
Identification info:

Description of element	Cycle										ΣT	$\bar{T}$
	1	2	3	4	5	6	7	8	9	10		
1												
2												
3												
4												
5												
6												
7												
⋮												
J												
											Select time	

Fig. 5-13 Sample time sheet.

jobs, etc. In this approach the job is sampled at random times until the pattern of the job has evolved. First we take a closer look at time study methods.

TIME STUDY

The first step in taking a time study is to carefully record, by means of a sketch, the physical characteristics of the work station, including the tools and fixtures, which will be considered to be the standard work practice of the job. Next the task is divided into its basic elements of motion. The elements should take at least 2 minutes to perform, as a finer breakdown is usually unwarranted. The elements are then recorded on a time sheet similar to the one shown in Fig. 5-13.

The worker who is selected to be studied should spend enough time on the task so that he knows the sequence well and performs it smoothly. With a stopwatch, he is timed for several cycles. The watch is usually kept going and the clock time at the completion of *each* element is recorded on the time sheet. At the end of the timing session each recorded time can be subtracted from the previous one so that clock time can be converted to actual time.

Select time Somewhere between 5 to 25 successive cycles are timed. When completed, the times recorded for each element are examined to see if any deviate substantially from the cluster of recorded values. If this does occur, the time can probably be identified with some unusual factor, such as the dropping of a tool, the breaking of a tool, material defect, etc. Since allowances for these factors will be made later, they should be eliminated from this first computation. Now the remaining cycle times for each element are added together. The results are placed in the ΣT column of Fig. 5-13. Next the average for each element, $\bar{T}$, is computed by dividing ΣT by the number of cycles which were observed. Finally these average element times are summed to give the average time it took to perform the entire task. This is called the *select time*.

Adjustment for pace Now, this select time represents the time it took a particular individual to perform a series of steps. Perhaps he is faster or slower than the average. Therefore, some adjustment should be made of this observed time so that it may reflect the pace of an average worker.

To accomplish this adjustment, the select time is multiplied by a rating factor, RF. In some cases this factor is determined subjectively. An estimate is made, at the time of the study, of the normal pace for the task. This is measured as the number of pieces that could have been completed by an average worker. The rating factor is then computed in the following way.

$$\text{RF} = \frac{\text{Actual Pace}}{\text{Subjective Estimate of Normal Pace}}$$

If the observed worker is considered to be slower than average, the RF will be less than 1. A rating factor greater than 1 indicates that the observed worker is faster than average. Some use experience alone to estimate a normal pace. Because of the subjectivity associated with this approach, many turn to a more objective method.

There is no perfectly objective method for rating a particular performance, but there are more effective means than the use of experience alone. Motion picture studies have been successfully used to train those responsible for establishing time standards. These movies illustrate a range of paces for a range of motions and train the observer to rate the subject quite accurately and consistently.

Allowances After the select time has been adjusted for pace, it must be adjusted for allowances. It is hardly reasonable to expect the worker to function every minute of a 480-minute day. Allowances should be made for personal time, fatigue, machinery problems, broken tools, and defective materials. The allowance for personal time includes breaks during the day, and these of course depend upon the kind of work being done. In very strenuous tasks and those which expose the worker to environmental difficulties such as extremes of temperature, breaks would come more frequently. Labor unions, as you would expect, often play a significant role in setting these allowances.

In some cases an allowance is made for fatigue. The argument in its favor points to the fall in performance as the day progresses. Those opposed to this allowance argue that if output rates are set properly in the first place, fatigue should not set in. All allowances considered, the total adjustment usually is in the vicinity of 20 percent.

Standard time Now a *standard time*, complete with all adjustments, can be calculated.

$$\begin{matrix}\text{Standard}\\ \text{time}\end{matrix} = \begin{matrix}\text{Select}\\ \text{time}\end{matrix} \times \text{RF} + \left[\text{Allowances}(\%) \times \begin{matrix}\text{Select}\\ \text{time}\end{matrix} \times \text{RF}\right]$$

This represents the average per-unit time it should take an average worker to continually perform a given task.

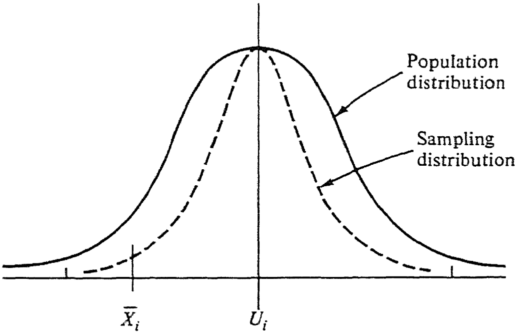


Fig. 5-14 Distribution of element times.

//// **The appropriate number of observed cycles** The process of observing a number of cycles and using the average of these observations as the estimate of the true but unknown actual average time for performing the task is nothing more than conventional statistical sampling. As in any sampling experiment, the first issue is the appropriate number of observations that must be taken. In time studies this becomes the appropriate number of cycles that should be timed. The number depends upon the confidence that is desired, the error that can be tolerated, and the variability of the dominating element in the task. We now turn to these problems.

Each element time can be looked upon as being generated by a process with mean U_i and standard deviation σ_i. An estimate of these parameters, $\bar{X}_i$ and S_i, is determined from a sample of n cycles. The sample mean $\bar{X}_i$, however, is only an estimate, and the probable magnitude of this error can be seen in Fig. 5-14. The population distribution is represented by the solid line, and the dotted line represents the distribution of sample means for samples of size n. The standard deviation of the sampling distribution is called the *standard error of the estimate* $\sigma_{\bar{x}}$. It can be approximated by

$$\begin{array}{c} S_i \\ n_i \end{array}$$

where S_i = standard deviation for that element computed from sample of n_i cycles

According to the central-limit theorem, the sampling distribution of the mean can be approximated by a normal distribution. For example, it can be said with a probability of .99 that $\bar{X}_i$ will differ from U_i by less than $2.56(S_i/\sqrt{n_i})$. This difference, $E_i = \bar{X}_i - U_i$, is sometimes called the *error of the estimate.*

We can also say with a probability of .95 that the error of the estimate will be less than some quantity E_i, where $E_i = 1.96(S_i/\sqrt{n})$. This relationship can be used to determine the sample size which is required to attain a certain degree of error and confidence. By solving for n_i we will find a sample size such that 95 percent of the time the maximum error of the estimate is E_i.

$$n_i = \left[\frac{(1.96)S_i}{E_i}\right]^2$$

This computation, however, must be undertaken for each element. Whichever element requires the largest sample can be said to dominate the others, and it will determine the number of cycles that need to be observed. Therefore

$$n^* = \max\,(n_1, n_2, \ldots, n_i, \ldots, n_j)$$

Unfortunately the variability of each element, S_i, is needed to compute the sample size, and it is seldom known before the sample is taken. Therefore, the usual procedure is first to sample between 5 and 25 cycles. From this, the sample variance of each element can be computed, and from it the sample size that *should* have been taken can be calculated. If the actual was greater than the computed size, the experiment is complete. If not, additional cycles need to be timed.

For example, if we take a preliminary sample for the first element of $n' = 20$ cycles and from it compute an $S_1 = 6$ seconds, the number of cycles that should have been observed in order to obtain a 95 percent chance that the error would be less than 3 seconds is

$$n_1 = \left[\frac{(1.96)(6)}{3}\right]^2 = 16$$

Computations are continued for the other elements that make up the task, and we find

$$\begin{aligned} n_2 &= 18 \\ n_3 &= 17 \\ n_4 &= 14 \\ n_5 &= 15 \\ n_6 &= 14 \\ n_7 &= 19 \\ \text{and}\qquad n^* &= 19 \end{aligned}$$

Since the actual sample of $n' = 20$ is greater than $n^* = 19$, the sample that has been taken is more than adequate to assure the error limitation and confidence required.

Rather than specifying an error for each element, it would be more convenient to specify the error in terms of a percent from the mean. For example, we could say that the sample size should be such that 95 percent of the time the error will be less than *5 percent from the mean.* We can modify our approach in the following way:

$$E_i = .05 \frac{\Sigma X_i}{n'}$$

$$1.96 \frac{S_i}{\sqrt{n}} = .05 \frac{\Sigma X_i}{n'}$$

$$n = \left(\frac{39n'S_i}{\Sigma X_i}\right)^2$$

where n = sample size that should be taken
n' = sample size that was taken
S_i = standard deviation of sample that was taken

Using the above formula, you will find the number of required observations for each element which will predict the element's time within 5 percent of its population mean 95 percent of the time. ////

PREDETERMINED TIME STANDARDS

The problem of adjusting observed times to an average pace can be avoided by the use of predetermined time standards. These standards have been carefully compiled and are available for a wide spectrum of motions. This, however, means that the job must be broken down into its basic elements, called *therbligs*. Then the time for each of these motions is located on the charts and recorded. After all the times have been added together, the appropriate allowance factor is added to generate the *normal time*. Since the many different kinds of users all have their own allowance factors, the predetermined time standards have been compiled without them.

The advantages of predetermined time standards are many. As already mentioned, it is not necessary to adjust the observed time to an average pace. Secondly, a worker does not have to be trained, and as a result the standards can be ready well before the production phase begins. The third advantage is that the predetermined standards have been well received by workers and their labor unions.

WORK SAMPLING

Work sampling is used where the cycle time is very long or where it is inconvenient, costly, or morale-destroying to initiate a full-scale time study. In its simplest application, it is used to determine the percentage of the time that a certain work activity is being performed. For example, in a large commercial bank a work sampling study was undertaken to determine the percentage of time that an office clerk spent filing information, typing, and answering the telephone. Using traditional work sampling methods, random observations were made of several clerks throughout the day. This process was continued for several weeks. The activity for each observation was recorded as either filing, typing, answering the telephone, or other. When the study was completed, the percent occurrences in each category was used to represent the true percent of the time that each of these activities occurs.

A more complex use of work sampling is to generate standards. To do this, however, it is necessary to know the total amount of work that was completed during the observation period. Also, the figures must be adjusted for the pace of the work being observed. Usually some overall average is taken.

SUMMARY

In this chapter we focused on the design problems associated with work content and work methods. It is here that we determine which steps must be included at the work station and the best possible method for executing these steps. Generally there is a considerable degree of flexibility in grouping steps, and the solution to this line-balancing problem is far from obvious. As a result, either optimization or heuristic techniques become welcome aids in the resolution of this problem.

An essential input to this stage is a set of labor standards. With them the process can be effectively designed, standards can be set, and the behavior of the system can be controlled. Since the days of Fredrick Taylor in the early 1920s, much attention has been focused on labor standards. Presently we have a sophisticated reservoir of knowledge and experience in this area; countless books have been written and courses offered on the subject.

One very important aspect of this whole process, however, is far from being solved. This is the role which human factors should play in the design phase. Frequently these factors overshadow technical ones, and the eventual solution is not one which was derived from a strict application of motion and time study. The human element deserves a chapter of its own. It is to this that we now turn.

QUESTIONS

Q5-1. What factors influence the choice of a cycle time?

Q5-2. In the assembly of an automobile, how much discretion is there in the selection of a cycle time, and why?

Q5-3. Is job specialization inevitable in some industries?

Q5-4. What is the objective in line balancing? How are cycle times set?

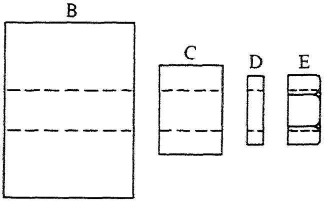
Q5-5. Suppose that you have been commissioned by the U.S. Post Office to determine the actual costs incurred by the different classes of mail. How could work sampling be employed to accomplish this task?

Q5-6. In a line-balancing problem with 10 steps and 6 precedent relationships, approximately how many feasible sequences could be established?

Q5-7. What are the advantages of predetermined time standards? Might their introduction into the plant system be difficult? Why?

Q5-8. What is an operation chart, and when is it used?

Q5-9. Construct an operation chart for the following assembly problem. Can your assembly method be improved by the incorporation of a fixture? If so, design one and construct the new operation chart.



Q5-10. Suppose that a new assembly technique has been suggested for one of the stations along a high-volume production line. What method of analysis would be appropriate for studying and perhaps improving upon this suggestion?

Q5-11. Why is the select time inadequate for representing the standard of job performance?

Q5-12. In the past, one of the automatic-lathe operators in a large machine shop has experienced a considerable amount of idle time while his machine was sequencing through its job. The foreman has suggested that this operator could run another lathe just like it. The machine operator, on the other hand, feels that this would not be economical. With two machines there would certainly be some machine idle time. On occasion this idle time could be quite substantial, since the setup and breakdown times for jobs vary considerably.

How would you resolve this problem?

Q5-13. What role would physiological factors play in the design of a work station where an operator must assemble 45 parts into a completed electronic subassembly?

Q5-14. How would labor standards differ between a retail sales clerk, a production-line employee, and a nurse?

Q5-15. Differentiate between time study standards and predetermined standards.

PROBLEMS

P5-1. Two stations along an assembly line have been balanced with an average of 60 seconds' work each. The arrival pattern for parts at the first station has been observed to be the following:

Time between arrivals

Time	*Frequency*
40 seconds	.10
50	.20
60	.40
70	.20
80	.10

The length of time that it will take the average operator at each station has been observed over several cycles, and it is also recorded below.

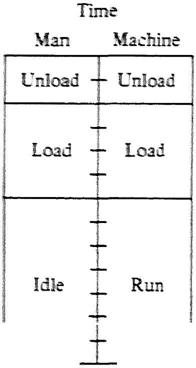
Cycle time

Time	*Frequency*
40 seconds	.15
50	.20
60	.30
70	.20
80	.15

Simulate the production line for a period of 15 minutes and estimate the number of parts per minute which can be expected to emerge from the second station. Assume no inventory between stations.

Can you tell from the results of your simulation if in-process inventories would help?

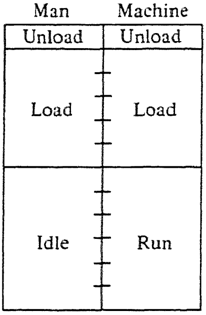
P5-2. Recently a forecast was made which shows that several additional punch presses must be purchased by the machine shop. At present, one machine is used to capacity and the machine operator does experience some idle time while the machine is running. A study of his cycle reveals that it takes 2 minutes to unload each piece, 4 minutes to load, and 7 minutes to run. This can be portrayed in the following way.



(*a*) If machine cost is $10 per hour and labor cost is $7 per hour, what is the cycle time, man idle time, machine idle time per machine, number of pieces produced per hour, and per-piece cost?

(*b*) Extend the analysis to one man, two machines and to one man, three machines. Which is the most economical?

P5-3. If an operator is assigned to a single machine, then his cycle can be depicted in the following way.



If machine cost per hour is \$5 and labor cost is \$10, how many machines should he operate to minimize per-piece costs?

P5-4. The following data have been collected from a time study of a new product:

(*a*) Number of cycles timed, 40
(*b*) Select time, 135 seconds
(*c*) Rating factor, .90
(*d*) Allowance factor, .25

What is the standard time to perform the job? Was the operator who was timed faster or slower than average?

P5-5. The following data have been collected from a time study of a new product:

(*a*) Number of cycles timed, 50
(*b*) Select time, 100 seconds
(*c*) Rating factor, 1.1
(*d*) Allowance factor, .15

What is the standard time to perform the job? Was the operator who was timed faster or slower than average?

//// **P5-6.** To determine the approximate number of cycles that should be observed, a preliminary sample of 30 cycles is observed. From it a standard deviation of $S_1 = 4$ seconds is computed for the first element in the task. How many cycles should be observed for this element in order to obtain a 99 percent chance that the error would be less than 2 seconds?

Computations for the other elements show the following number of cycles which should be observed:

$n_2 = 30$ $n_5 = 22$
$n_3 = 17$ $n_6 = 18$
$n_4 = 21$

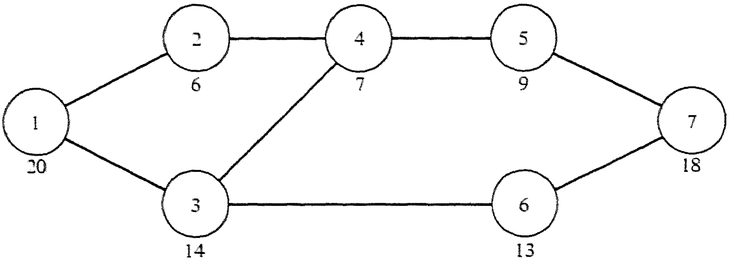
What strategy should be chosen? ////

//// **P5-7.** How many cycles should be observed for element 1 when a preliminary set of 20 observations resulted in a standard deviation of 8 seconds? The criterion for sample size is that there be a 95 percent chance for the error to be less than 4 seconds. ////

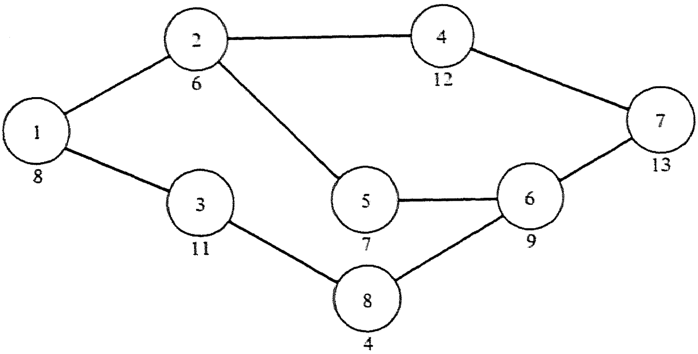
//// **P5-8.** What sample size should be taken for a particular element so that 99 percent of the time the error will be less than 6 percent from the mean? Derive the formula.////

//// **P5-9.** Using the formula which you developed in Prob. P5-8, find the number of cycles which should be timed for the first element when a preliminary sample of 20 cycles resulted in a mean of 80 seconds and a standard deviation of 5 seconds. ////

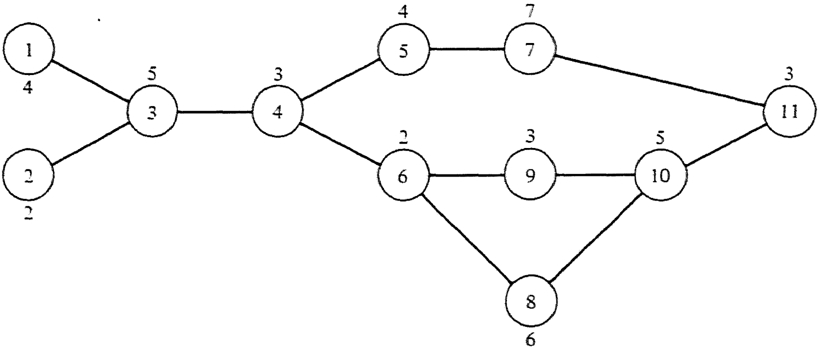
P5-10. The following is a set of steps which must be divided among a group of work stations in such a way as to minimize idle time. Assume that cycle time is set at 30 seconds. Set up a linear programming model of this problem. Do not solve.



P5-11. Set up the following line-balancing problem in linear programming format. Let the cycle time be 20 seconds.

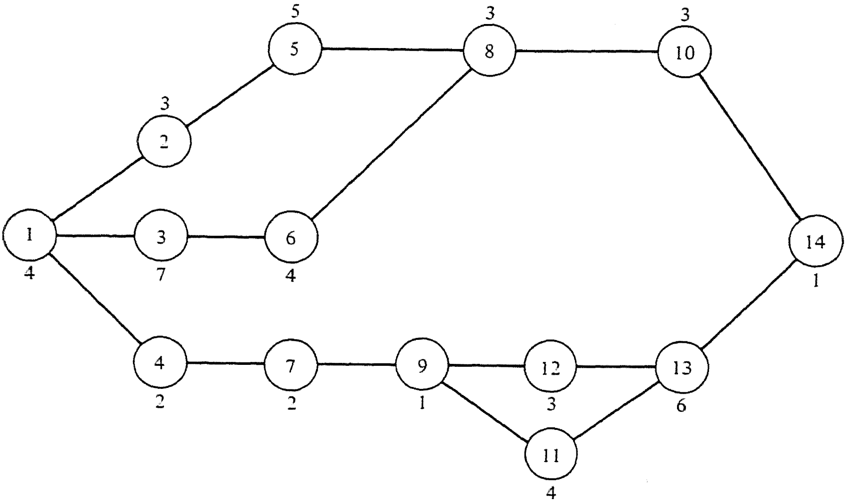


P5-12. Solve the following line-balancing problem, using the heuristic approach. Let the cycle time be 10 minutes.



P5-13. Solve Prob. P5-12 for a cycle time of 8 minutes.

P5-14. Using the heuristic model, determine the minimum number of stations which are needed and the steps to be included at each of these stations for the following line-balancing problem. Let the cycle time be 10 minutes.



P5-15. Solve Prob. P5-12 for a cycle time of 14 minutes.

BIBLIOGRAPHY

Barnes, Ralph M.: *Motion and Time Study: Design and Measurement of Work*, 6th ed., John Wiley & Sons, Inc., New York, 1968.

———: *Work Sampling*, 2d ed., John Wiley & Sons, Inc., New York, 1957.

Buffa, Elwood S.: *Modern Production Management*, John Wiley & Sons, Inc., New York, 1969.

Ignall, Edward J.: "A Review of Assembly Line Balancing," *Journal of Industrial Engineering*, July–August 1965.

Kilbridge, M. D., and L. Wester: "A Heuristic Method of Assembly Line Balancing," *Journal of Industrial Engineering*, vol. 12, no. 4, July–August 1961, pp. 292–298.

Mastor, Anthony A.: "An Experimental Investigation and Comparative Evaluation of Production Line Balancing Techniques," *Management Science*, vol. 16, no. 11, July 1970, p. 728.

6

Human Factors in Job Design

INTRODUCTION

The previous chapters, which focused on the macro and micro design characteristics of a transformation process, treated the design of the system from a mechanistic point of view. Underlying this analytic approach was the assumption that there is a "best way" to design the process. Once this "best way" is uncovered, the optimal operation of this man/machine system is assured. It was therefore expected that the individual would adapt to the role necessitated by the job. Little, if any, attention was directed toward the individual who must fill this role.

The purpose of this chapter is to review and analyze the research that has been undertaken to bridge this gap between the requirements of the job and the needs of the individual. Five main stages serve to identify the development of these ideas. The first is the engineering approach associated with Taylor's scientific management school, the second is represented by the human relations school, the third by the job enlargement school, the fourth by the job participation school, and finally that embodied in the job enrichment school.

ENGINEERING DESIGN

If anything can categorize the scientific management school, it is the search for the "one best way" to undertake a job. To accomplish this, two dimensions were emphasized: first, the coordination of the physical transformation process; secondly, the coordination of the human being into this physical process. The focus on the individual, however, was strictly task-oriented. Psychological and sociological factors which could not be linked immediately to productive output were of no concern.

To improve the efficiency of the physical transformation process, several steps were taken. First, the practice whereby the worker would choose his own technique was abandoned. Standard manufacturing practices were developed. Secondly, standards of performance, which had been heretofore worker-determined, were also developed. Time study was the method for setting these standards.

Then, to coordinate the human being with this physical process, wage incentives based on a piece rate system were introduced. The support of the worker was won when it became clear that it was possible to earn higher wages under this plan than under the traditional one. Under this new system the worker benefited by taking home more money and the company benefited by a substantial increase in productivity.

The upshot of the scientific management movement was a mechanistic view of the transformation process as well as a mechanistic view of the worker's role in that system. As a result much more control of the worker's environment was transferred to management, and along with this transfer went the loss of several sources of worker satisfaction.

HUMAN RELATIONS

The human relations school took as its domain the psychological and sociological environment of the job. Its framework evolved from a series of studies undertaken at the Hawthorne plant of the Western Electric Company. The initial objective of this study was to determine the effect that certain variables, which were associated with the physical environment, had on output. These variables included light intensity, temperature, humidity, and hours of sleep. In all of these experiments the transformation process was treated as a given.

The major finding from these experiments was that changes in the physical environment did not suffice to explain changes in the rate of output; the researchers had to look elsewhere. Focus then shifted to the social group, and it was successfully determined that the social environment contributed to these changes in output. Performance of the group was related to satisfaction for the individual, cohesion within the work group, coordination between the work group and supervision, and the collective purpose of the group. The policy that emerged from this school made a revolutionary impact on management style; it was: strengthen the work group, build cohesion within the group, and harmonize the plant.

In its popular version, "human relations" to many practitioners meant that a happy group was a productive group. True, a happy group probably exhibits strong cohesion, but there is no evidence to show that a cohesive group will motivate its members toward task-oriented goals. In fact, quite frequently cohesive groups can be found repressing the achievement of task-oriented goals. The congruency between the group goals and task-oriented goals is essential to predicting performance levels of cohesive groups. If the group's goals, as defined by their members, are congruent with the task, performance levels will be high. If there is low congruence, performance will be low.

The astute practitioner of human relations therefore showed concern for the needs of the group *and* the requirements of the task. His job was one of integration and balance. The consequence of this new focus was a rejection of the mechanistic point of view taken by scientific management and the introduction of the social aspects in the work environment.

JOB ENLARGEMENT

According to the human relations school, performance was strongly tied to the social forces acting on individuals. The content of the job was considered to be technologically determined and outside the bounds of discretionary design. They accepted the "one best way" philosophy of the scientific management school.

The problem that had evolved from continued application of these scientific management principles was that many jobs seemed to have become specialized, routine, mindless, boring, and unsatisfying. The job enlargement school therefore reasoned that since much of the satisfaction was derived from the job itself, perhaps satisfaction and performance could be improved by its redesign.

JOB DEPTH AND SCOPE

To deal effectively with job enlargement, two aspects of the job must be distinguished: job scope and job depth.[1] If the job contains but few steps and the cycle is repeated frequently, the *scope* is said to be narrow. Typical of a job with a narrow scope is the assembly job found in the automated production lines of automobile manufacturers. An example of a job with a broad scope is that of a mailman. His job cycle takes one day.

Job depth refers to the control one has over his working environment. If the worker has no influence in planning his work and he is dependent on decisions made by his superiors, job depth is said to be low. The assembly-line worker who is dominated by his paced line can show little independence in fulfilling his work role. He therefore has a job which is low in depth and narrow in scope. A research physicist, on the other hand, has a job which is high in depth, since he can exercise considerable control over his working environment. He can show considerable independence in his work. The physicist, then, has a job which is high in depth *and* broad in scope. The postman's job, although broad in scope, is relatively low in depth—that is, he has little control over his working environment.

---

[1] Allan C. Filley and Robert J. House, *Managerial Process and Organizational Behavior*, Scott Foresman and Company, Glenview, Ill., 1969, p. 214.

The objective of job enlargement is to increase both the scope and the depth of the job. For example, the scope of the assembly-line worker's task could be enlarged by assigning more steps, thereby increasing the cycle time. The depth could be enlarged by removing the job from a paced line and in addition making the worker responsible for the quality of his work. The latter could be accomplished by having the worker stamp his assembly with his mark.

There are, of course, technological and economic limitations which affect the degree of job enlargement. It would hardly seem reasonable to have a worker build a Chevrolet from scratch. Certainly it would be a satisfying experience for him, but what would such a strategy forbode for our standard of living? Therefore a return to a craftsmanlike ideology is probably an impossibility, given our desires for mass consumption of material goods which only a highly technology-oriented economy can provide to its consumers. There is a constraint imposed on the extent of job enlargement, but the constraint differs between industries and between jobs within these industries. In the high-volume, highly capital-intensive industries such as the automotive, chemical processing, and steel industries, the constraints are very binding. In others, such as the electronics industry, hospitals, public administration, and most all services, the constraints are not nearly as limiting during the job design phase.

CASE STUDIES

Of the several job enlargement studies, those undertaken by the Maytag Company and IBM stand out. As a result of their job enlargement program, Maytag reported an improvement in quality, lower labor costs, and a preference for the job enlargement environment by most of the operators. These conclusions were drawn after a 6-year period in which 25 job enlargement projects were undertaken. Maytag felt that legitimate conclusions could be drawn since 16 of these job enlargement projects were previously conveyor-paced group assemblies. Of the 130 employees who worked on job enlargement projects, 90 worked on both types.

Perhaps the flavor of job enlargement can best be illustrated by an example of job enlargement. Maytag produced a water pump which was used in one of their automatic washers. This pump contained 26 components, which before job enlargement were assembled by a group of from five to seven operators, including one man for relief, repair, and stockup. Cycle time was from .33 to .40 minute per unit, depending upon the number of workers in the group. All steps were carried out while the assembly was fixed to a moving conveyor.

Under a program of job enlargement, the job was redesigned so that one operator completely assembled a pump. The cycle time was increased from an average of .33 minute to 3.0 minutes per unit. In addition, the operator was also responsible for the quality of his work. Therefore the scope and depth of the job were increased.

As a result of this program, management reported that quality had improved, operator turnover was substantially reduced, and costs were reduced. Direct labor costs decreased some 16 percent while indirect costs decreased 50 percent.

In the IBM study, job depth and scope were also increased. Before the change, four separate jobs, including machine operator, setup man, tool sharpener, and inspector, existed. Under a job enlargement program these jobs were combined into one. The machine operator assumed full responsibility for his machine and its output. The consequence of this was increased worker satisfaction, lower production costs, and better quality.

In its popular version job enlargement is misinterpreted just as was the human relations school. To some practitioners job enlargement means increasing the scope of the job. It is therefore proposed, by them, that if a job requiring two steps is enlarged to include five steps, the worker's satisfaction will increase. Such results seem unlikely. It is unreasonable to expect that three mindless and boring steps added to two mindless and boring steps will lead to job satisfaction. It is *both* the *scope* and the *depth* that require analysis in any job enlargement program. In fact, when you look beneath the surface of Maytag's program, it becomes evident that even more than these two factors account for their successful results.

The company had, for some time, fostered a management style that took into consideration both the needs of the individual and the demands of the job. Even before their job enlargement program, attitude surveys were taken and suggestion programs were firmly established. In keeping with this management style, a participative environment was structured for each job enlargement undertaking wherein employees became involved in the development of the new work design. For example, one trial bench was constructed so that the operators could have a chance to work under these new conditions, evaluate the job, and make suggestions for improvements.

SCIENTIFIC RIGOR

Unfortunately the Maytag and IBM job enlargement programs were not performed with a high intensity of research rigor. For example, a *field experiment* approach would have been more fruitful. In a field experiment the variables which are expected to be affected by the experiment are identified and defined. In addition, the procedure for collecting data on these variables is carefully recorded. Next, one or several testable hypotheses are developed. Finally, a series of experiments designed to test these hypotheses is executed. Many times, a control group and experimental group are utilized in the field study. The experimental group would be subjected to the change or stimulus being studied while the control group would not. Results comparing these two groups more adequately describe the response to experimental changes, since other changes in the environment not under the control of the experiment are the same for both groups.

What is the main advantage of a scientifically controlled field experiment? If the experiment is carefully described, the variables of interest identified and defined, hypotheses stated, and results carefully recorded, it is possible for others to replicate the study. For it is only after several replications in which findings corroborate the initial evidence that a management principle can be said to be established.

The Maytag study was not a field experiment. Some of the concepts, such as the increase in job satisfaction that was achieved by the workers, were very tenuously defined. In addition, data and explicit results were not reported. Much of the analysis was retrospective; the study and its outcome were recorded well after many of the programs were undertaken. Consequently much of the detail was lost to the passage of time.

Many studies have focused on job enlargement, and yet the evidence is still unclear, there is no unanimity of opinion that it is a good thing. Confusion in the comparison of one study with another is the rule, not the exception. This results because the consequences of job enlargement are multidimensional. It is reported to affect quality, productivity, and satisfaction. Often, in fact, it is not clear how these variables should be measured. Satisfaction, for example, has been measured by absenteeism, turnover, attitude toward job, and attitude toward supervisor, to name a few. Most studies therefore focus on different variables, measure them differently, attain different levels of research rigor, and weigh the outcome of these variables differently in making general statements about job enlargement.

JOB ENLARGEMENT AND SATISFACTION

Many job enlargement studies have focused on job satisfaction with the underlying belief that if satisfaction is increased, motivation and performance will follow. However, both satisfied and unsatisfied workers have been found in specialized and enlarged jobs. In one study of a knitwear mill it was concluded that feelings of monotony were not only a consequence of the task performed but were related to the worker himself. The study further concluded that workers who were content while performing repetitive tasks displayed such personal characteristics as contentment with present state of affairs, placidity, and possibly rigidity; they accepted things the way they were. These conclusions suggest that job specialization is not the root of the problem. Perhaps individual differences account for some of the variance in experimental outcomes.

Kennedy and O'Neill[2] studied the effect of job content on workers' satisfaction at a large automotive assembly plant. They utilized *correlational* techniques in which observations are taken on two or more variables for a sample of individuals or groups. The direction and degree of the relationships that exist are found, using statistical techniques. No attempt is made by the investigator to manipulate any of these variables.

---

[2] James E. Kennedy and Harry E. O'Neill, "Job Content and Workers' Opinions," *Journal of Applied Psychology*, vol. 42, no. 6, 1958, pp. 372–375.

Kennedy and O'Neill administered a 71-item questionnaire designed to uncover any difference between the satisfaction of assembly-line workers and that of utility men. The cycle time for the assembly workers was between 1 and 2 minutes; the task was routine and repetitive. The utility men, on the other hand, had many functions. They relieved assembly operators for scheduled or emergency breaks and served as instructors to new operators, corrected production errors, and sat in for absent operators. Comparison of statistical results showed no difference in worker satisfaction between the two jobs.

Walker and Guest,[3] on the other hand, found evidence to the contrary. The degree to which employees in an automobile assembly plant found satisfaction in their job was directly related to the number of steps which composed their job.

The evidence on job enlargement is therefore quite mixed. More than likely, a singular focus on the content of the job is not enough to explain differences in personal satisfaction and differences in attainment of task-oriented goals such as productivity and quality. Job enlargement by itself might, then, be considered an oversimplified solution to a complex problem.

JOB PARTICIPATION

One problem identified in the Maytag study was the likely development of resistance to change upon introduction of a job enlargement program. Management's sensitivity toward this problem led them to create an environment of participation in the hope of overcoming this resistance. Lack of scientific rigor in their program prevented the association of this variable with their results.

COCH AND FRENCH STUDY

Coch and French[4] highlighted the importance of job participation through several field experiments in which a high level of scientific rigor was employed. The experiments were carried out at Harwood Manufacturing Company, a pajama manufacturer.

The employees at Harwood were on an individual piece rate system where 60 units of work per hour was regarded as the standard efficiency rating. Pay was tied to weekly efficiency, and if the worker attained an average efficiency rating of 70 for the week, he would receive a 10 percent bonus. An efficiency rating of 80 would be accompanied by a 20 percent bonus, and so on.

If the job content was changed, or if the worker was shifted to another job, a transfer bonus was given. This would ensure that no loss in earnings would be suffered during the learning process. In spite of this, attitudes toward job changes were quite negative at Harwood. Resistance was expressed as grievances toward new piece rates, high turnover, low efficiency, restriction of output, and aggression toward management.

---

[3] C. R. Walker and R. H. Guest, *The Man on the Assembly Line,* Harvard University Press, Cambridge, Mass., 1952.

[4] Lester Coch and John R. P. French, Jr., "Overcoming Resistance to Change," *Human Relations,* 1948, vol. 1, pp. 512–533.

An analysis of several hundred experienced operators who had undergone job changes illustrates this resistance to change. Only 38 percent of them recovered to the standard efficiency rating of 60 units. The remainder either quit or became permanently substandard operators. The learning rate for those who recovered is illustrated in Fig. 6-1. Of some surprise is the fact that the learning rate for new operators was faster than that for transfers. This led to the conclusion that skill is of secondary importance in the rate of recovery, whereas motivation is of primary importance.

To overcome this resistance to change, the focus was directed toward the work group. A field experiment was designed in which four such groups were designated. The first group was directed to appoint several group representatives. In the second and third groups all members participated, and the fourth acted as a control group.

The control group was subjected to their usual routine for a job change. A group meeting was held where it was explained that the need for change was the consequence of competitive pressures. A new piece rate was set, questions were answered, and the group was dismissed.

In group 1, the approach was quite different. The operators were called into a meeting and shown a competitive garment which sold for much less money than their product. Through this exposure the group started to share some of management's needs: steps had to be taken to reduce the price of the product and remain competitive. It was agreed that certain "frills" and "fancy" features could be eliminated from the product without affecting the worker's chance to

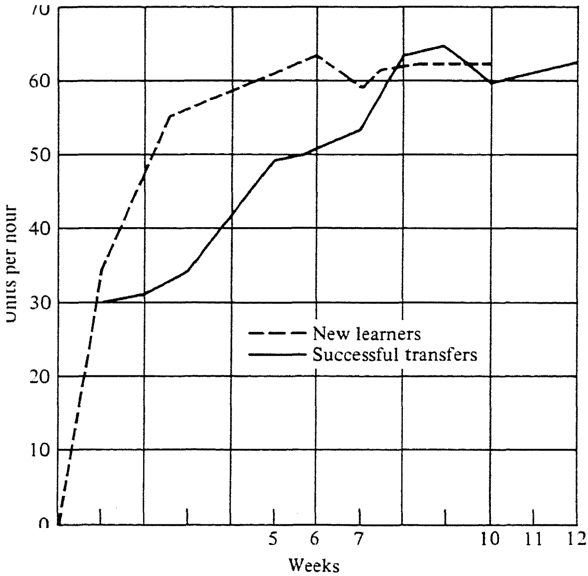


Fig. 6-1 A comparison of the learning curve for new, inexperienced employees with the relearning curve for only those transfers (38 percent) who eventually recover to standard production. (*From Lester Coch and John R. D. French, Jr., "Overcoming Resistance to Change," Human Relations, 1948, 1, p. 514.*)

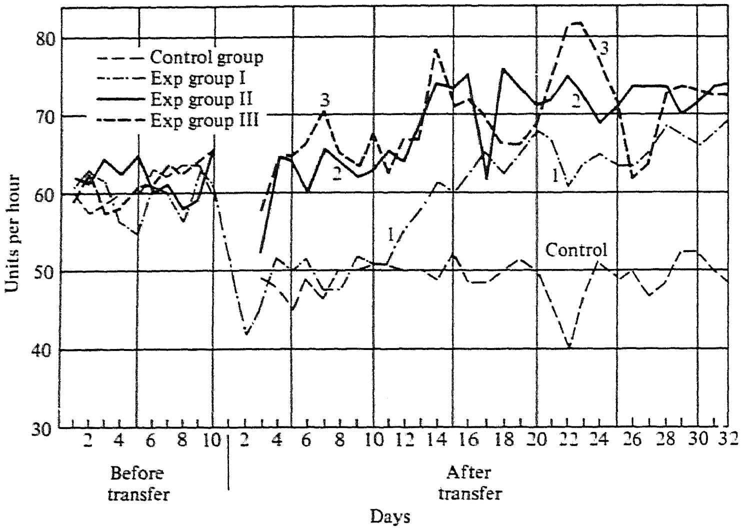


Fig. 6-2 The effects of participation through representation (group 1) and of total participation (groups 2 and 3) on recovery. (*From Lester Coch and John R. D. French, Jr., "Overcoming Resistance to Change," Human Relations, 1948, 1, p. 522.*)

reach full efficiency. A few members of the group were appointed to study the necessary changes. At a later meeting they made several suggestions concerning the new jobs. They were then trained under these new methods and later acted as instructors and trained the rest of the group.

Groups 3 and 4 went through the same kind of process with the following exception: all members of the group participated instead of just a select few.

The results of these experiments are summarized in Fig. 6-2. It can be seen that the control group improved little. The usual conflicts between the group and management erupted. Seventeen percent quits were recorded in the first 40 days. Group 1 showed very good learning rates, and very little conflict was recorded. Groups 2 and 3 recovered much faster than group 1. They eventually recovered to an efficiency rate 14 percent higher than the pre-change level. Furthermore, no conflicts were recorded.

The conclusion reached by the study was that representative participation shows good results, but full participation shows even better results. What about other studies? Vroom, after analyzing the conclusions of several field experiments and correlational studies, concluded the following:

> When the entire pattern of results is considered, we find substantial basis for the belief that participation in decision making increases productivity. There is experimental and correlational evidence indicating that higher levels of influence by workers in making decisions that they are to carry out results in higher productivity than lower levels of influence.[5]

[5] Victor H. Vroom, *Work and Motivation*, John Wiley & Sons, Inc., New York, 1964, p. 226.

JOB ENRICHMENT

Industrial psychologists rather than focusing on one of these schools of thought prefer to analyze both the needs of the individual and the requirements of the job and then determine the extent of harmony or conflict between them.

HIERARCHY OF NEEDS

The writings of Professor Abraham Maslow represent the classic works which delve into the needs of man. According to his theory, man's needs are arranged in a hierarchy. Higher-order needs emerge only after the lower-level ones are satisfied. The lowest-level needs, which must be satisfied first, are physiological. They include man's need for food, clothing, and shelter. Once these needs are satisfied, attention shifts to the next level, safety. Following safety in the hierarchy we find love, esteem, and, finally, self-actualization. The highest needs of man are, then, the needs to be self-actualized. This represents the search for man to be what he is.

The school of thought to which we now turn returns the emphasis to the content of the job, for it feels that only through intrinsic involvement in one's own work can a person achieve the highest of all needs, self-actualization. All other factors in the environment, such as supervision, participation, and monetary incentives, are seen as playing a secondary role. Motivation, they feel, can be generated only by the role incumbent—it cannot be imposed externally. This theory is known among practitioners as the *job enrichment school*. Its roots can be traced to Herzberg's two-factor theory of job satisfaction.[6]

TWO-FACTOR THEORY

Herzberg's initial study of job satisfaction was limited to 200 engineers and accountants and was criticized when results were generalized. However, more than 20 replications of this study over a wide range of populations has been undertaken since then, making this one of the most replicated studies in the field.

His hypothesis was that factors which produce job satisfaction are separate and distinct from those which lead to job dissatisfaction. Since two sets of factors are involved here, satisfaction and dissatisfaction are not opposite kinds of feelings. That is, the opposite of job satisfaction is *not* job dissatisfaction but no satisfaction, whereas the opposite of job dissatisfaction is not job satisfaction but no job dissatisfaction, as shown in the following diagram:

No satisfaction	Motivators	Satisfaction
Dissatisfaction	Hygiene	No dissatisfaction

Herzberg suggests that two kinds of needs are involved here: the lower-level needs of man, which are stimulated by the job environment, and the higher-level needs of man, which are stimulated by the job content. When the lower-level needs are not met, dissatisfaction occurs. Herzberg refers to these environmental factors as *hygiene factors.* These stimuli include company policy and administration supervision, interpersonal relationships, working conditions, salary, status, and security.

---

[6] Frederick Herzberg, *Work and the Nature of Man*, World Publishing Company, Cleveland, 1966.

Higher-level needs are stimulated by factors intrinsic to the job. Herzberg calls these stimuli *motivator* factors. They include achievement, recognition for achievement, the work itself, responsibility, and growth or advancement.

Herzberg compiled the results of 12 correlational studies in which some 1,685 respondents were asked to describe sequences of events when they were exceptionally satisfied or dissatisfied with their job.[7] These studies covered such occupational groups as lower-level supervisors, professional women, agricultural administrators, men about to retire from management positions, hospital maintenance personnel, manufacturing supervisors, nurses, food handlers, military officers, engineers, scientists, housekeepers, teachers, technicians, female assemblers, accountants, Finnish foremen, and Hungarian engineers.

Analyzing the responses to this question, it became apparent that dissatisfaction revolved around aspects in the job environment, while satisfying experiences revolved around aspects intrinsic to the job. The results supported the hypothesis.

Figure 6-3 summarizes the results of these 12 studies. It can be seen that motivators were mentioned much more frequently when some experience led to job satisfaction. The insert at the lower right of Fig. 6-3 shows that, on an overall basis, 81 percent of the factors contributing to job satisfaction were motivators. Hygiene factors, on the other hand, were mentioned much more frequently when an experience was extremely dissatisfying. In all, 69 percent of the factors contributing to employee dissatisfaction could be traced to hygiene factors.

The predominant job characteristics which led to satisfactory experiences included achievement, recognition, work itself, responsibility, and advancement. An interesting and important job (work itself), job responsibility, and advancement were found to be the most important factors that could lead to a lasting change.

The characteristics of the job that led to dissatisfaction included the company policy and administration, with supervision as the most important. Interpersonal relations with supervisors, work conditions, and salary led less frequently to dissatisfaction. Salary, however, seems to have been mentioned equally as a satisfier and as a dissatisfier.

JOB ENRICHMENT STUDY

The two-factor theory therefore suggests that the way to achieve worker satisfaction is to enrich the content of the task. In general it could be said that the focus should be more on the depth than on the scope of the task. Several

---

[7] Frederick Herzberg, "One More Time: How Do You Motivate Employees?" *Harvard Business Review,* vol. 46, no. 1, 1968, p. 53.

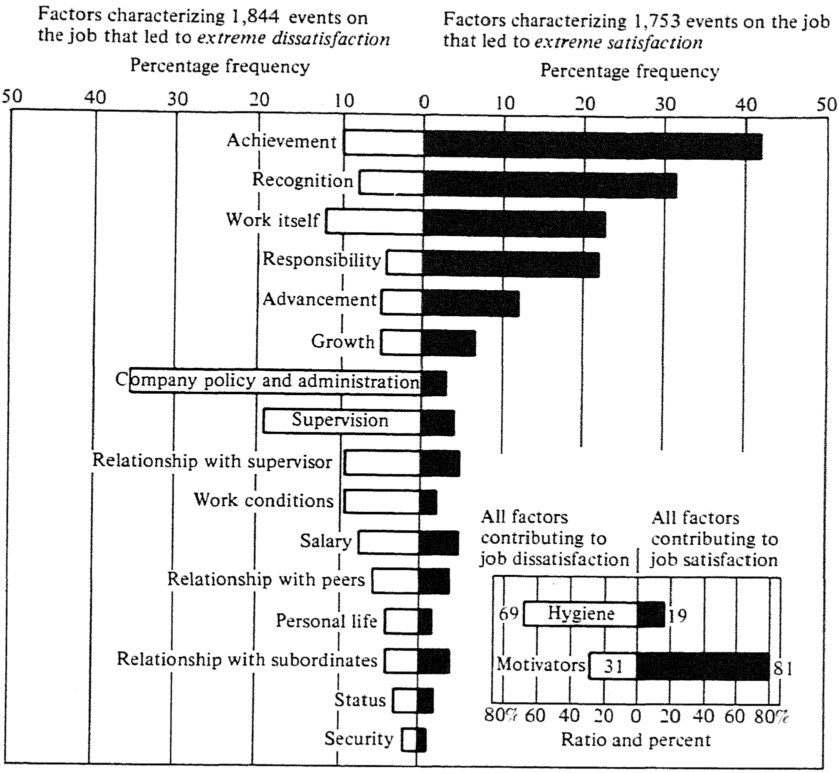


Fig. 6-3 Job characteristics and job satisfaction. (*From Frederick Herzberg, "One More Time: How Do You Motivate Employees?" Harvard Business Review, vol. 46, no. 1, 1968, p. 57.*)

principles which can be used to guide a job enrichment effort have been suggested by Herzberg and are presented in Table 6-1.

Herzberg also conducted a field experiment designed to verify the basic prediction of his theory, that job enrichment will lead to an increase in job satisfaction and worker productivity. The subjects of this study were highly educated women who functioned as stockholder correspondents for a large corporation. Prior to the introduction of the job enrichment program, almost all indexes of performance and job attitudes were low; the job was no challenge.

The field experiment consisted of one experimental group and one control group. The control group merely continued to do its job in the usual way. The experimental group, on the other hand, underwent some significant changes. The nature of their jobs was enriched by the introduction of changes based on the principles summarized in Table 6-1. The specific changes are summarized in Table 6-2.

The results were in the direction expected. Job satisfaction was significantly improved for the experimental group while the control group actually

Table 6-1 Principles of job enrichment

Principle	Motivators involved
1. Removing some controls while retaining accountability for the task	Responsibility and personal achievement
2. Increasing the accountability of the individual for his own work	Responsibility and recognition
3. Giving a person a complete unit of work	Responsibility, achievement, and recognition
4. Granting additional authority to an employee in his job. More control over his environment	Responsibility, achievement, and recognition
5. Making periodic reports directly available to the worker himself rather than to his supervisor	Internal recognition
6. The introduction of new and more difficult tasks	Growth and learning
7. Assigning individuals specific or more specialized tasks which will enable them to become experts	Responsibility, growth, and advancement

From: Frederick Herzberg, "One More Time: How Do You Motivate Employees?" *Harvard Business Review*, vol. 46, no. 1, 1968, p. 59.

Table 6-2 Job changes in field experiment

1. Individuals within the group were appointed as experts in certain subject areas. They were consulted with before a worker would seek supervisory help.
2. The workers signed their own names to the letters. Previously the supervisor had done all the signing.
3. Proofreading by a supervisor was dropped from 100 percent to 10 percent. Previously every letter was proofread.
4. Whereas pressures to produce were previously imposed on the group, production was discussed only in terms of "a full day's work is expected."
5. Outgoing mail now went directly to the mailroom without first going to the supervisor's desk.
6. The form letter was abandoned. Workers were encouraged to respond in a more personal way.
7. Responsibility for the quality and accuracy of letters was now vested in the worker and not the supervisor.

From: Frederick Herzberg, "One More Time: How Do You Motivate Employees?" *Harvard Business Review*, vol. 46, no. 1, 1968, p. 61.

suffered a decrease in satisfaction. The magnitude of this change can be seen in Fig. 6-4. Performance for the experimental group also improved. It was measured by the quality of letters, including accuracy of information, and speed of response to stockholders' letters of inquiry. The results are shown in Fig. 6-5. From the graph it can be seen that the experimental group was performing lower than the control group before the experiment began but soon reversed their position.

INDIVIDUAL DIFFERENCES

The two-factor theory is by no means unanimously supported. Not all replications have supported this theory. This could be attributed to the omission

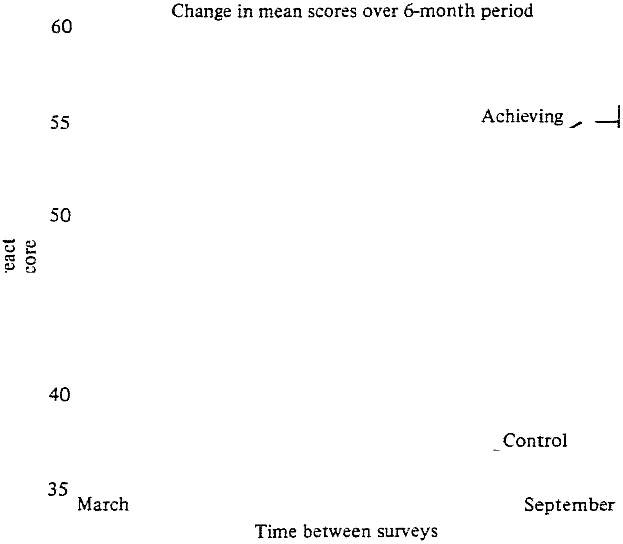


Fig. 6-4 Attitudes toward tasks. (*From Frederick Herzberg, Work and the Nature of Man, World Publishing Company, Cleveland, 1966, p. 162.*)

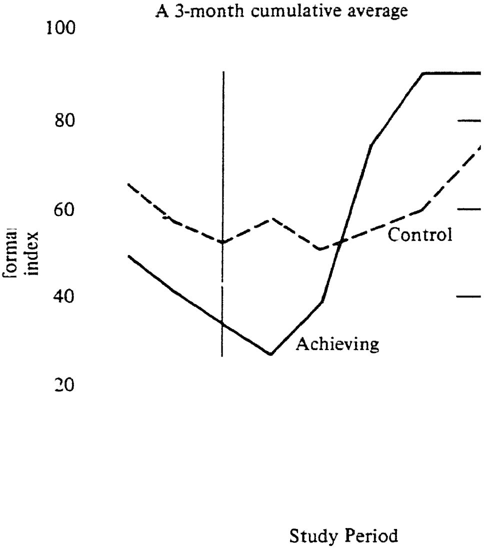


Fig. 6-5 Performance index. (*From Frederick Herzberg, Work and the Nature of Man, World Publishing Company, Cleveland,* 1966, *p. 163.*)

of a very important variable: individual differences. Several studies have suggested that a certain job characteristic can cause job satisfaction for one person and job dissatisfaction for another. It can be concluded therefore that there is substantial variance associated with the degree of satisfaction achieved by individuals from a job characteristic. It is unreasonable to expect that satisfaction can be explained or predicted from the job characteristic alone or from the personality characteristics alone. Rather, both should be considered simultaneously.

Variables which can help predict whether a given job characteristic will be a source of satisfaction or dissatisfaction include the occupational level of the job and the individual's age, sex, educational attainment, culture, and standing in his group. Despite the fact that this school is more scientifically based than the job enlargement school, omission of individual differences has led some to accuse the two-factor theory as also being guilty of oversimplification.

JOB SATISFACTION, MOTIVATION, AND JOB PERFORMANCE

Underlying many of these schools of thought is the implication that job satisfaction is directly related to the motivation to perform; if job satisfaction increases, performance levels will rise. This hypothesis is based on the assumption that as the worker becomes more and more satisfied, he will reward his employer by accepting and achieving management's goal of high productivity.

Contrary to this common assumption, research indicates that there is no significant relationship between job attitudes and job performance.[8] It also shows that those characteristics which lead to job satisfaction may not be identical to those characteristics which motivate the individual and lead to job performance. In fact, some situations may exhibit low satisfaction and high performance, others high satisfaction and high performance, and still others low performance. It may be that we have to treat each situation as a separate outcome of the person-work role relationship.[9]

The reason that there appears to be no direct link between satisfaction and performance is because of the intermediate role that motivation plays. For only when properly motivated will even the satisfied individual meet performance objectives.

Motivation is related to the degree of congruency between the needs of the individual and the demands of the job. If the needs of the job are highly congruent with the needs of the individual, a high level of motivation to perform effectively is likely to occur. If the needs and demands are negatively related, the individual can be said to be negatively motivated and a low level of performance will likely result. The more positive the relationship, the higher the level of motivation to perform effectively.

Performance, however, is not strictly dependent on the level of motivation. It is also a function of the individual's ability. The relationship can be illustrated in the following way:[10]

$$\text{Performance} = f(\text{Ability} \times \text{Motivation})$$

---

[8] Arthur H. Brayfield and Walter H. Crockett, "Employee Attitudes and Employee Performance," *Psychological Bulletin*, vol. 52, no. 5, 1955, pp. 396–422.

[9] Vroom, op. cit., p. 187.

[10] Ibid., p. 203.

Since performance is a function of the product of ability and motivation, the hypothesis leads to the conclusion that for low values of ability, increases in motivation will lead to only small increase in performance. For high levels of ability, however, increases in motivation will lead to large increases in performance.

Motivation to perform effectively is therefore a complex result of job characteristics and personality characteristics. Studies of motivation based solely on the characteristics of the work role or solely on personality characteristics have generally been found to be inadequate. There is, however, some valuable insight to be gained by reviewing these studies. It is to them that we now turn.

SUPERVISION AND PERFORMANCE

Job characteristics that have been studied include supervision, the work group, job content, wages, and promotional opportunities. Some of these operate in the same direction for motivation and satisfaction, some operate in opposite directions, and for others the direction is not at all clear. First we turn to studies which have focused on the relationship between supervisors and motivation.

Several studies have shown that consideration by a supervisor for the needs and feelings of his subordinates is positively related to the motivation that they display in performing their tasks. There can be situations, however, where this relationship does not hold: Would not military relationships between officers and enlisted men be an example of this?

Even though consideration and performance have been shown to be related, there is still some question as to the direction of causality. Does consideration lead to job performance, or does a high-performing group end up winning the consideration of their supervisor? The only way to uncover the direction of this relationship is by *laboratory experimentation.* In a laboratory setting the consequence of different levels of consideration on individual performance could be examined. Other environmental factors would be held constant. The problem in performing this kind of experiment would be in changing supervisory behavior. In spite of the absence of conclusive evidence, it seems reasonable to conclude that consideration tends to affect job performance.

Supervisory influence on job performance can also be affected by the degree of job participation encouraged by the supervisor. There is both field experiment and correlational evidence showing that higher levels of participation and influence lead to higher productivity, while lower levels lead to lower productivity.

The second situational factor that can affect job performance is the work group. The results of research, including the Hawthorne study, indicate that groups which are highly cohesive exert more of an influence over their members than groups low in cohesion. The direction that this cohesiveness can take, however, is not predictable. Cohesiveness can have either a negative or a positive influence over individual performance, depending upon the group performance norm. Job performance will be strong where the individual strongly desires acceptance by the group and anticipates acceptance only through effective performance. Job performance will be weak, on the other hand, where group cohesion is again strong but acceptance is anticipated only if performance is weak. A study by Patchen[11] illustrates the impact of the supervisor's role on group performance norms. The following factors were considered important in explaining this impact. First, the supervisor should have high *standards of efficiency* for the group. It is reasonable to expect employees to think that if performance matters little to the supervisor, then rewards and punishment will not be based on it. Secondly, the supervisor should have the *power* to deliver rewards and punishment, if his efficiency standards are to matter. He must therefore have some control over pay increases, promotions, vacations, and work pace. Thirdly, the employees should be convinced that he is willing to use his power effectively. That is, he must be willing to "go to bat" for the employees if the occasion does arise.

The results showed that none of these characteristics when taken by themselves are strongly related to performance norms. When combinations of these characteristics were compared, however, some strong relationships were uncovered. For example, among those foremen who encouraged efficiency, going to bat for employees by trying to obtain rewards for them is strongly associated with high group norms.[12] However, among foremen who do not encourage efficiency, going to bat for employees has a negative effect on group performance norms. If we divide the supervisors according to the extent that they go to bat for their subordinates, we find that for those supervisors who often go to bat for their men, encouraging efficiency brings an increase in group performance norms. Supervisors who go to bat for their men much less frequently, however, achieve the opposite effect if they encourage efficiency. These results are summarized in Table 6-3. Therefore an action like going to bat for employees can have different outcomes depending on whether the supervisor does or does not encourage efficiency. The point that begins to emerge from this research, then, is that several aspects of a supervisor's behavior must be "considered if his impact on group performance norms is to be understood."

---

[11] Martin Patchen, "Supervisory Methods and Group Performance Norms," *Administrative Science Quarterly,* vol. 7, 1962, pp. 275–293.

[12] $r = .86$.

Table 6-3 Supervisory methods and group performance norms

	Encourages efficiency	*Goes to bat for men*
Goes to bat much	positive	—
Goes to bat little	negative	—
Encourages efficiency much	—	strong positive
Encourages efficiency less	—	negative

What Patchen considers as the most significant finding of this study is that while close supervision, by itself, does not lead to high group performance norms, if it is associated with a supervisor who is a rewarding figure, together with a high level of group cohesiveness, then high group performance norms follow. The definition of "close supervision" in this study refers to frequent checking on subordinate's work and not on reducing the worker's freedom to perform the task in his own way. When all of these factors are operative, an environment is created where the cohesive group accepts the supervisor as a member of their team and as one who is genuinely interested in their welfare. As a result the interactions between supervisor and subordinate are many and an apparent pattern of close supervision emerges.

JOB CONTENT AND PERFORMANCE

Job content can also affect motivation, but less attention has been given to the motivational aspects of this characteristic than to its association with job satisfaction. That specialization can increase or decrease productivity is clear; the question that must be answered is: Under what conditions will one or the other of these results occur?

WAGES AND PERFORMANCE

Wages that are directly tied to performance have been shown to be significantly related to the level of performance. Thus workers on an individual work incentive plan would be more motivated by this characteristic than would hourly paid clerks. This relationship is strongest for those individuals whose need for economic remuneration is high.

Promotional opportunities are also related to the degree of motivation. If the individual believes that promotion is tied to job performance, then a strong level of motivation will exist. If, on the other hand, the individual perceives little relationship between performance and promotion, he will perform less effectively. If the individual is indifferent toward promotion, it is not one of his needs and promotional opportunities would not be related to his performance.

SUMMARY

As you proceeded through this chapter, it should have become apparent to you that there are no easy guidelines for effectively designing a working environment that meets both the needs of the individual and the demands of the job. Certainly we can say that to design the job with only the demands of the job taken into consideration is to ignore "the human side of enterprise," a position that is becoming increasingly untenable in a modern society.

In the scientific management school, the perspective was limited to the lower-level needs of man. Money was seen as the means by which these needs could be satisfied. But this was only the beginning.

As psychologists increased their understanding of man's needs, and as the standard of living increased, thereby satisfying these lower-level needs, the focus began shifting to the higher-level needs.

The speed with which this shift can be realized depends in part upon technological progress, for as more jobs are automated, fewer menial, routine, and boring jobs remain. The movement is certainly in this direction, since the percentage of the labor force employed in a factory environment has been steadily decreasing in recent history. As this trend continues, the technological constraints that prevent the higher-level needs from being considered in the job design will not be nearly as binding. Evidence of this can be seen in the larger number of people being employed by service industries, hospitals, and public administration agencies. It can therefore be concluded that the opportunity to design meaningful and rewarding jobs is increasing.

The effectiveness of job design depends upon the needs of the individual who will fill that role and the demands of the job. We found that the human relations, job enlargement, job participation, and job enrichment schools each focused on but a few dimensions of a multidimensioned problem. The job enrichment school was probably the most comprehensive, yet by not considering individual differences it has been accused by many theoreticians of being an oversimplified theory. From a practical point of view, however, these theories do provide valuable guidance in the design of jobs.

Rather than focusing on a few dimensions, industrial psychologists try to uncover and verify those factors that can best explain changes in job satisfaction and job performance. To do this, they must study both job characteristics and individual characteristics, as well as the complex interaction between them. Their study methods are both scientifically designed and scientifically executed and take one of three forms: field experiment correlation study, and laboratory experiment.

Since it is unlikely that each job could be designed with a particular individual in mind, job characteristics might be the most useful aspect of the psychologists' findings. Many of the important job characteristics that affect satisfaction and performance, such as the group process, the work itself, and participation, were the focus of the earlier schools. Therefore we have much supportive evidence that points to their importance in the job design effort.

QUESTIONS

Q6-1. Is it necessarily true that a happy work group is also productive?

Q6-2. What are the differences between the way in which the scientific management school and the job enrichment school perceive the worker?

Q6-3. On what source of motivation did the scientific management school rely? Do you think this source is relevant today?

Q6-4. Does the increase in the number of steps included at a work station necessarily lead to the desirable effects associated with job enlargement?

Q6-5. What is participative management? Does the evidence show that it can be helpful in reaching output objectives? Why would it affect worker satisfaction?

Q6-6. Might individual differences account for the success or failure of a job enlargement program?

Q6-7. Contrast job enlargement with job enrichment.

Q6-8. What is the two-factor theory of job motivation? How would you make use of it when designing the following jobs?

(*a*) Nurse
(*b*) Assembly-line worker
(*c*) Postman
(*d*) Salesgirl
(*e*) Inventory clerk
(*f*) Secretary
(*g*) Bank teller
(*h*) Laboratory technician

Q6-9. How is job satisfaction linked to job performance?

Q6-10. The HP Company manufactures several types of electronic instruments. One of these, an oscillator, has been in production for some 20 years. The process engineers have recently decided that the assembly process should be given a fresh look. In the past the instrument has been assembled by assembly-line methods. Each person performed an average of 10 steps and then passed the unit to the next station. The cycle time was approximately 3 minutes.

One of the process engineers has suggested that each worker should assemble a complete instrument. In all, 120 steps are necessary to complete a unit. This means that it will take each worker approximately 40 minutes to complete his cycle. At present demand levels, 10 workers would be continuously supplied with work.

A major problem that must be overcome before this new system can be implemented is the design of the work station to store efficiently the 75 parts that are used in the assembly process. The process engineer who has suggested the new method believes that he has the solution. It is a two-tier Lazy Susan system which will not only hold all of the parts but will swing them into the front position in their desired sequence.

All that remains now is to make a decision between the two methods. Which method would you prefer? Why? What are the pros and cons behind your choice?

BIBLIOGRAPHY

Filley, Allan C., and Robert J. House: *Managerial Process and Organizational Behavior,* Scott, Foresman and Company, Glenview, Ill., 1969.

Herzberg, Frederick: "One More Time: How Do You Motivate Employees?" *Harvard Business Review,* vol. 46, no. 1, 1968, p. 53.

———: *Work and the Nature of Man,* World Publishing Company, Cleveland, 1966.

Vroom, Victor H.: *Work and Motivation,* John Wiley & Sons, Inc., New York, 1964.

7

Project Design

INTRODUCTION

In this chapter the problems associated with the planning of a project will be considered. A *project* is defined as a task of considerable magnitude involving both time and money that is usually, but not always, carried out but once. Examples of this include the building of a hospital wing, the construction activities associated with a world's fair, a research and development program for an ethical drug, the withdrawal of troops from Vietnam, the construction of an elementary school, the introduction of a new product, the design, manufacture, and deployment of a missile system, the repair and construction of ships, and the installation of a computer system.

These examples all share certain characteristics. First, they all have a starting and a finishing point. Second, some activities can be carried out independently of others. Some, moreover, can be initiated only upon completion of earlier activities. For example, in building construction, the roof can be started only after framing is complete. Third, each activity requires the expenditure of that scarce resource, time.

Because of these characteristics, a plan which focuses on the scheduling of such tasks should be developed before the first activity is allowed to proceed. If the project contains but a few activities, an informal approach can be taken. However, in complex projects it is necessary to develop a formal plan if some degree of efficiency is desired in its execution.

In addition to scheduling the sequence of activities, the plan can be used to control the progress of the project. For it should become clear from the plan which activities are critical to the completion of the project. If these critical activities are delayed, the scheduled completion of the whole project is jeopardized. These critical activities should therefore be carefully monitored.

In this chapter we investigate several formalized approaches to the project scheduling problem. The most widely known include PERT and CPM. PERT, which is an acronym for "program evaluation and review technique," was developed as a planning tool for the Polaris missile project. The problem was first to design a realistic schedule for the countless number of activities that were required to design, develop, test, and manufacture the missile system, and then to use the schedule as a control device after the project had begun.

CPM is an abbreviation for "critical path method." It was developed independently of PERT by the Du Pont Company. There are, in fact, some major differences between these two methods. PERT takes into consideration the uncertainty that surrounds the estimate of completion time required for an activity. For example, in a missile system where a new alloy must be developed, it would be unrealistic to expect the head of a research and development lab to estimate that this activity will take exactly 5 weeks. A more realistic answer might convey his uncertainty in an explicit manner. For example, he might feel that if everything goes well, it could be done in as short a time as 3 weeks, with average luck 5 weeks, and with terrible luck 9 weeks. This uncertainty is taken into consideration by the PERT technique. Shortly you will see exactly how this is accomplished.

CPM is used more for projects where the uncertainty surrounding the time estimates can be considered insignificant. So rather than focusing on uncertainty, CPM is concerned with the trade-off between cost and the completion date of a project. It examines the consequences of utilizing more manpower to shorten the project completion time.

NETWORK DESIGN

Now we turn to the way in which we can formally structure these projects as a network of activities. Our first step is to be a little more explicit about some of the terms we have been using.

ACTIVITIES

Activities are the basic jobs, tasks, or operations that must be performed. Occasionally an activity can be initiated completely independently of other activities. That is, it can be performed at any time during the project cycle. The usual case is that an activity has one or several precedent relationships. It is therefore essential that these precedent relationships be carefully established when the project is decomposed into its basic activities. Precisely how far decomposition should go is a difficult question to answer. It depends upon the purpose of the planning activity and on the ability to separate activities. For example, in some cases it might be appropriate to identify the "installation of a plumbing system" as an activity whereas in others it might be more appropriate to decompose this into the 100 detailed steps which are undertaken, from the drawing of the plumbing specifications to the final test and inspection of the installed system.

EVENTS

An *event* marks the beginning or the end of an activity. Frequently we will use an event to mark the end of one or more activities *and* the beginning of one or more. Therefore the major difference between activities and events is that activities represent the passage of time while events are special points in time.

NETWORK

Now we are ready to combine the concepts of activities, events, and precedence into a network. Consider the example summarized in Table 7-1. A hospital wishes to renovate a section of their main building and equip it with modern cardiac monitoring equipment. First the administrators must forecast the demand for service of this nature. This is, in turn, followed by the design of the new wing, selection of the equipment, construction, and finally the installation and checkout of the equipment.

Some of these activities must be preceded by others. From Table 7-1 we can see that the forecast phase, or activity *a*, is the first step in the project and as such has no predecessor. The design of the wing and the selection of the electronic equipment, however, must both be preceded by the completion of the forecast. Construction must be preceded by design, and finally, the installation and checkout phase must be preceded by both the completion of construction and equipment selection.

These relationships can be illustrated on a network diagram. In Fig. 7-1 the activities are shown as arrows and the circles represent events. For example, event 1 represents the start of the forecasting phase, and arrow *a* represents the duration of this phase. Event 2 represents the completion of the forecasting

Table 7-1

Activity	*Alternate designation*	*Description*	*Immediate predecessor*	*Time*
a	1,2	Forecast demand	—	5 weeks
b	2,3	Design new wing	*a*	6
c	2,4	Selection of equipment	*a*	5
d	3,4	Construction	*b*	12
e	4,5	Installation and checkout	*c*,*d*	4

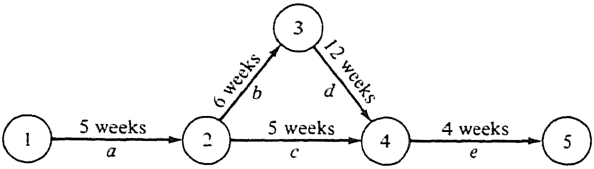


Fig. 7-1 Project network.

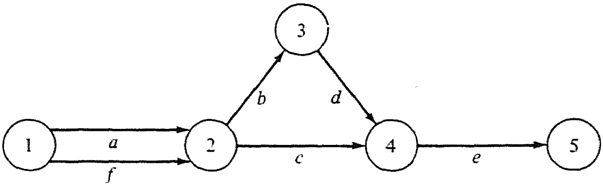


Fig. 7-2 Erroneous implication that activity *f* must precede activity *c* as well as activity *b*.

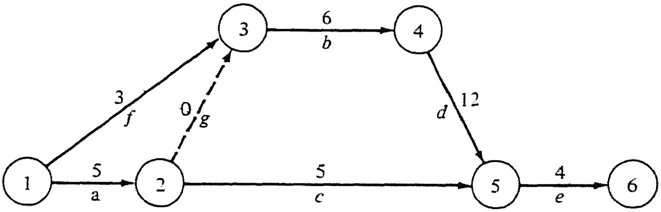


Fig. 7-3 Introduction of dummy activity.

phase *and* the initiation of both the design phase and the equipment selection phase. The estimated time necessary to complete each activity is entered above the arrow.

DUMMY ACTIVITIES

Consider the addition of an activity which involves the selection of an architect and general contractor. This activity can take place simultaneously with forecasting but must precede *only* the design phase. We will call it activity *f* and estimate that it takes 3 weeks to complete. If, however, we represent this activity by arrow *f* in Fig. 7-2, we imply that activity *f* must precede activity *c* as well as activity *b*. To overcome this problem, a dummy activity must be introduced.

A dummy activity is used only to indicate precedence and requires zero time to perform. In Fig. 7-3 the dummy activity *g* is introduced which makes activity *a* a predecessor of activities *b* and *c* but makes activity *f* a predecessor of only activity *b*.

We can generalize from this example that dummy activities are required when two or more activities have some, *but not all*, of their immediate predecessors in common. When this occurs, all of the predecessors should be drawn leading to separate nodes and dummy activities should be added from these nodes to comply with the precedence constraints. Since activities *a* and *f* were the immediate predecessors in our example, they are drawn leading to separate nodes. Dummy activity *g* was then added to make activity *a* become a predecessor to *b*.

CRITICAL PATH

Of primary interest in project scheduling is the minimum length of time that completion of the project will take. To determine this, it is necessary to find the longest path through the network. This path is called the *critical path*.

Consider the network shown in Fig. 7-1. There are two different paths through the network, the first one being 1-2-3-4-5 and the second one 1-2-4-5. The first path takes $5 + 6 + 12 + 4 = 27$ weeks, while the second path takes $5 + 5 + 4 = 14$ weeks. It can therefore be concluded that the total project will require 27 weeks for completion.

The path 1-2-3-4-5 is considered the critical path because completion of this path takes the longest. Each job on the critical path, moreover, is called a *critical activity*. If a critical activity is delayed, the length of time needed to complete the project will increase. For example, if activity *b* is delayed 2 weeks the project time increases from 27 weeks to 29 weeks.

When jobs do not fall on the critical path, some delay will have no effect on the completion date of the project. For example, activity *c* does not lie on the critical path, and as a result it can undergo some delay before it will start affecting the completion date. In fact, it can suffer a delay of 13 weeks before this can happen. We therefore conclude that activity *c* has 13 weeks of *slack*. The concept of slack and critical activities is a very useful one to management, for it points out which activities must be watched closely and in which ones there is some flexibility. We now turn to a set of formalized procedures for finding critical paths and slack activities in more complex networks.

Early start and early finish times Consider the example shown in Fig. 7-4. The expected time for completion of an activity is given above the arrow and to the left of the bracket. If we define the *early start* of an activity as the earliest

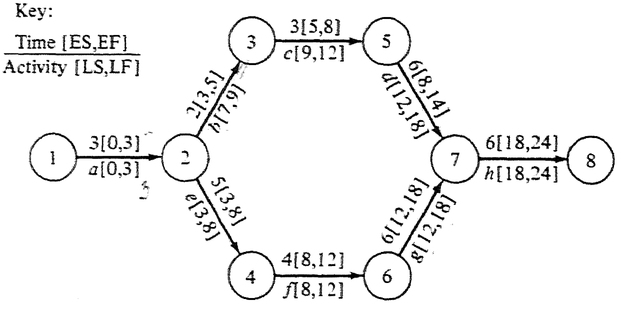


Fig. 7-4 Critical path network with early start, late start, early finish, and late finish times.

possible time that it can begin, the earliest start for activity *a* is zero. We enter the early start in the left-hand side of the top bracket. The earliest finish time is equal to the earliest start plus the expected completion time for that activity. The earliest finish time for activity *a* is therefore 3 weeks. "3" is entered in the right-hand side of the top bracket.

Before activity *b* can begin, its immediate predecessor, *a*, must be completed. Therefore, the earliest that *b* can begin is 3 weeks. Since it requires 2 weeks, the earliest that it can be completed is 5 weeks after the start of the project. The earliest start and finish times are calculated in a similar manner for activities *c*, *d*, *e*, *f*, and *g*. The earliest start for activity *h*, however, requires that we compare the earliest finish times for all activities which merge into this node. Activities *d* and *g* are immediate predecessors to activity *h*, and as a result activity *h* can commence only when *both* are complete. Since *d* is complete at the beginning of week 14 and *g* is not complete until the beginning of week 18, activity *h* cannot commence until week 18. In general, then, the highest number merging into the node will specify the earliest start time.

If the earliest completion time is 24 weeks, the longest path through the network is 24 weeks. This path, 1-2-4-6-7-8, is the critical path.

Late start and late finish time We discovered that activities which were not on the critical path could be delayed without affecting the earliest finish date of the project. Now we will make quite explicit the maximum delay which could be tolerated. This will be accomplished by specifying the latest time an activity can be started and the latest time it can be finished. Violation of either will mean project delay.

We proceed by working from the end of the network to the beginning. Starting at node 8, then, the latest that this activity can be finished is 24. This is entered in the right-hand side of the lower bracket. Since the activity takes 6 weeks, the latest that it can be started is $24 - 6 = 18$. The latest start is entered in the left-hand side of the lower bracket. There are two activities which merge at node 7. The latest that either could finish is 18. Otherwise the project would be delayed. The latest start for activity *d* is $18 - 6 = 12$. Computations for the earliest and latest start times for activities *b*, *c*, *e*, *f*, and *g* are entered into the lower brackets of Fig. 7-4. Two activities merge from activity *a*; therefore, in working backward to determine *a*'s latest finish, the *smallest* of the two latest start times is chosen. For example, if activity *a* is completed later than week 3, the project will be delayed.

Slack As we examine Fig. 7-4, we find that the early start and late start for some activities are identical. For these activities there is no flexibility; if the start is delayed, the project will be delayed. We can therefore conclude that critical activities are those for which the early start and late start are identical. In addition, the string of critical activities 1-2-4-6-7-8 make up the critical path. Those activities, however, for which the difference between the late and early start is positive have slack. For example, the early start for activity b is 3 and the late start is 7; therefore as much as a 4-day delay can be incurred before the project is delayed. We can conclude that slack can be computed in the following way:

$$S(b) = LS(b) - ES(b) = 7 - 3 = 4 \text{ days}$$

It is important to note that if we use some of these 6 days of slack in completing activity a, the computed slack for activities which follow may be affected. Finally it should be observed that if the latest start of a noncritical path is violated, a new critical path is born.

//// LINEAR PROGRAMMING FORMULATION OF NETWORKS

The simple deterministic project network which was covered in this section can be expressed as a linear programming model. First we let X_i represent the early occurrence of event i, where $i = 1, 2, 3, \ldots, k$. The objective is to minimize the early occurrence of the last event k. We can therefore write the objective function as

$$\text{Minimize } X_k$$

Constraints are necessary to ensure that the assigned time for an activity is at least as great as the required time. Therefore if we let t_{ij} be the required time for completion of the activity between events i and j,

$$X_j - X_i \geq t_{ij} \qquad \text{for all activities } ij$$

Finally we must specify the nonnegativity constraints.

$$X_i\text{'s} \geq 0$$

Consider the example shown in Fig. 7-5. The linear programming problem can be written in the following way:

$$\begin{aligned}
&\text{Minimize } X_5\\
&\text{Subject to}\\
&X_2 - X_1 \geq 3\\
&X_4 - X_1 \geq 10\\
&X_3 - X_1 \geq 7\\
&X_4 - X_2 \geq 5\\
&X_5 - X_4 \geq 3\\
&X_5 - X_3 \geq 7\\
&X_1, X_2, X_3, X_4, X_5 \geq 0
\end{aligned}$$

Using a standard LP code, the solution to this problem was found to be $X_5 = 14$. Therefore the earliest that the project can be completed is week 14.

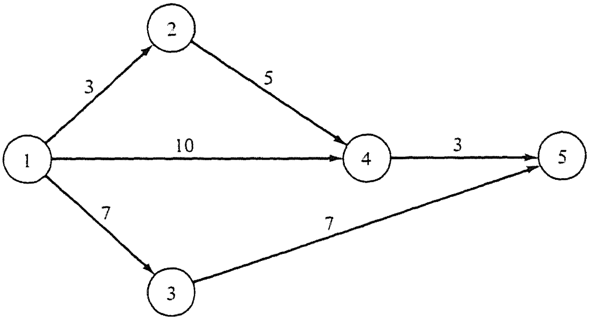


Fig. 7-5 Linear programming example.

The remaining events take on the following values:

$$X_1 = 0 \qquad X_3 = 7 \qquad X_5 = 14$$
$$X_2 = 6 \qquad X_4 = 11$$

There are, however, several alternative solutions which will still yield the same value for X_5. This occurs because of slack in activities 1-2, 2-4, 1-4, and 4-5. Therefore, X_2 and X_4 could be changed without affecting X_5.

To ensure that the solution reflects the earliest possible start time for each event, the objective function must be changed to read

$$\text{Minimize} \sum_{i=1}^{k} X_i$$

The constraints go unchanged. ////

PERT

Up to this point we have assumed that each activity time was known with certainty. In many cases this becomes an oversimplification that cannot realistically be accepted. In the case of a research and development project it would seem unrealistic to assume that we could know exactly how long each stage of the project would take. For in these kinds of projects, uncertainty is a natural phenomenon. There are countless projects where uncertainty is significant, and to ignore it would result in a superficial analysis. We now turn to a method for incorporating this added dimension into our analysis, PERT.

EXPECTED TIME

In PERT it is assumed that the administrator responsible for each activity can make three estimates, the first being the *most likely* time that it will take for the activity to be completed and the second being an *optimistic* time—a time which would result if everything went exceptionally well. Finally he is asked for a *pessimistic time,* a time which would result if the activity ran into difficulty.

The computation of the *expected time* for each activity depends on these three estimates. They are combined in the following way:

$$t_e(i) = t_o(i) + 4t_m(i) + t_p(i) \tag{7-1}$$

where $t_e(i)$ = expected time for activity i
$t_o(i)$ = optimistic time for activity i
$t_m(i)$ = most likely time for activity i
$t_p(i)$ = pessimistic time for activity i

The expected time for an activity represents the average time it would take if the activity were performed over and over again. Although this is not quite representative of the actual situation, it is usually the best estimate possible.

Consider the following example. As part of a research and development project, an electronic circuit must be designed, built, and tested. The engineer in charge feels that the most likely time for completion is 8 weeks. However, with unusually good luck that time could be reduced to as little as 5 weeks. With bad luck, however, it could take as much as 12. Underlying these three estimates is a continuum of possible times between t_o and t_p, with t_e being its arithmetic mean and t_m the mode. This is illustrated in Fig. 7-6. Since it might be very difficult to obtain the exact nature of this probability distribution, these three estimates are considered satisfactory in estimating its mean and standard deviation. Using Eq. 7-1, the mean is computed in the following way:

$$t_e = \frac{5 + 4(8) + 12}{6} = 8.2$$

Next we turn to the problem of estimating the standard deviation of the activity.

If the activities that make up the project display some variability, this source of uncertainty should be incorporated in statements concerning the project's length. One such measure of variability, the standard deviation of each activity, can be estimated by taking one-sixth the difference between the optimistic and the pessimistic estimates.

$$S_i = t_p(i) - t_o(i)$$

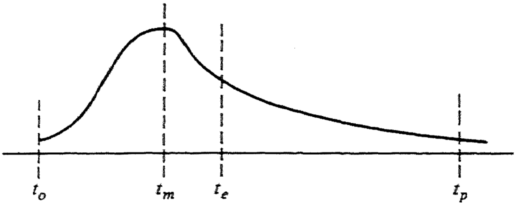


Fig. 7-6 Distribution of PERT time estimates.

In our example

$$S_i = \frac{12 - 5}{6} = \frac{7}{6} = 1.16$$

Since variance is equal to the standard deviation squared, we have

$$V_i = (\tfrac{7}{6})^2 = \tfrac{49}{36}$$

Now we turn to an example where these concepts are utilized in constructing and analyzing a PERT network.

THE CRITICAL PATH OF A PERT NETWORK

Consider the research and development project whose network is presented in Fig. 7-7. The expected time, standard deviation, and variance for each activity are shown in Table 7-2. The network is redrawn in Fig. 7-8; however, this time the expected time t_e is entered above the activity arrow. The critical path is 1-3-4-5-7, and the expected completion time is $8 + 7 + 5 + 13 = 33$ weeks. The variance of the critical path is equal to the sum of the variances associated with every critical activity along the critical path. Therefore, $V = 4 + 4 + 1 + 9 = 18$.

PROBABILITY ESTIMATES FOR COMPLETION DATES

We have estimated that the expected completion time is 33 weeks. However, since it has been acknowledged that some variability does exist for each activity,

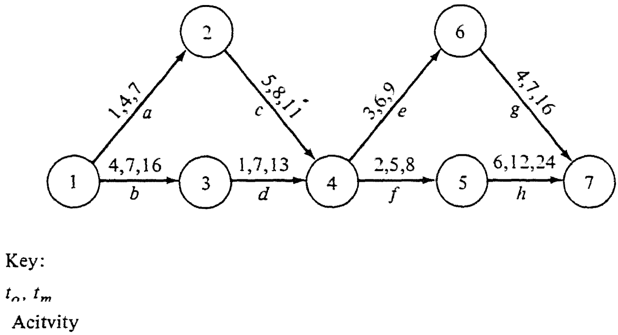


Fig. 7-7 PERT network.

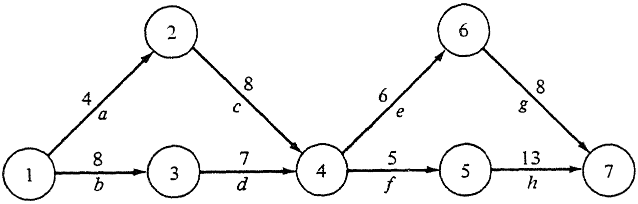


Fig. 7-8 Expected times for PERT network.

Table 7-2 Computation of expected time, standard deviation, and variance for PERT analysis

Activity	*Expected time*	*Standard deviation*	*Variance*
a	4	1	1
b	8	2	4
c	8	1	1
d	7	2	4
e	6	1	1
f	5	1	1
g	8	2	4
h	13	3	9

other completion times on either side of 33 weeks are possible. The likelihood of other times depends upon the variability computed for the critical path. According to the central-limit theorem of statistics, it is reasonable to assume that the variability associated with any path through the network, including the critical path, can be approximated by a normal distribution. Therefore in our example the completion time of the project can now be represented as a normal distribution with a mean of 33 and a standard deviation of $\sqrt{18} = 4.2$.

As a result of its normality, it is possible to make probability estimates of completion dates by referring to a normal table. For example, we can say that there is a 68 percent chance that the project will be complete within plus or minus 1 standard deviation of its mean. This stems from the fact that 68 percent of the area under a normal curve is within 1 standard deviation of its mean. In our example, it therefore follows that there is a 68 percent chance that the project will be completed between 28.8 and 37.2 days after starting. There is a 95 percent chance that the project will be completed within 2 standard deviations of the mean, or between 24.6 and 41.4 days, and there is a 99 percent chance that the project will be completed within 2.56 standard deviations of the mean, or between 22.3 and 43.7 days. These probability statements are illustrated in Fig. 7-9.

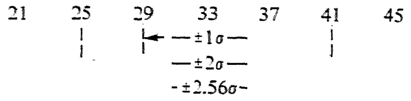


Fig. 7-9 Probability distribution of completion dates.

SOME ASSUMPTIONS BEHIND THE ANALYSIS

Behind our analysis are two assumptions that should be considered when using this approach. The first is that we have an unbiased estimate of the expected times for each activity. This means that if we used the expected time over and over again for our estimate, on the average we would be right. The second assumption is that the activity times are independent of one another. For example, this assumption implies that activity *a* is independent of activity *c*, and that even if activity *a* is behind schedule and there is pressure on activity *c*, activity *c* will not respond by a completion time earlier than expected.

THE CONSEQUENCE OF NEAR-CRITICAL PATHS

Not always is it safe to estimate a probability range for project completion based solely on the mean and variance of the critical path. Completion of the project may also hinge on the mean and variability of near-critical paths. If there is a noncritical path, for example, whose length is close to the critical path and whose variance is larger, it is certainly possible for this path to become critical. Therefore, using uncertainty analysis, which focuses only on the critical path, would result in an erroneous probability statement. If we used conventional statistical techniques to take this added dimension into consideration, the analysis would prove to be quite cumbersome. A much more manageable approach is to resort to simulation.

SIMULATION OF A PERT NETWORK

The first step in the simulation process is to draw the project network. In Fig. 7-10 a network is shown for the construction of a special-purpose steam turbine. Next we must determine subjective probability distributions for the completion times of each activity. For example, it was determined that the engineer in charge of activity *a* felt that there was a 70 percent chance that his activity could be finished in 5 weeks, a 10 percent chance of completion in 4 weeks, a 5 percent chance of 3 weeks, a 10 percent chance of 6 weeks, and a 5 percent chance of 7 weeks. This and the remaining subjective probability distributions are shown in Fig. 7-10.

The object of each replication will be to generate a project length based on its critical path. After this process is repeated a number of times, a distribution can be drawn depicting the length of the project.

The first step, therefore, is to generate an activity time for each activity, using random numbers. Second, the critical path for this trial is recorded, as is the completion time for this path. Now steps 1 and 2 are repeated several times. The results of 25 replications are summarized in Table 7-3.

The average completion date is 35.3 days. A distribution of the completion dates as shown in Fig. 7-11 conveys much more information than the simple average; from it we can estimate the likelihood of different completion times. For example, we can conclude that the likelihood that the project will take more than 37 days is 12 percent. What is the likelihood that it will take more than 35 days?

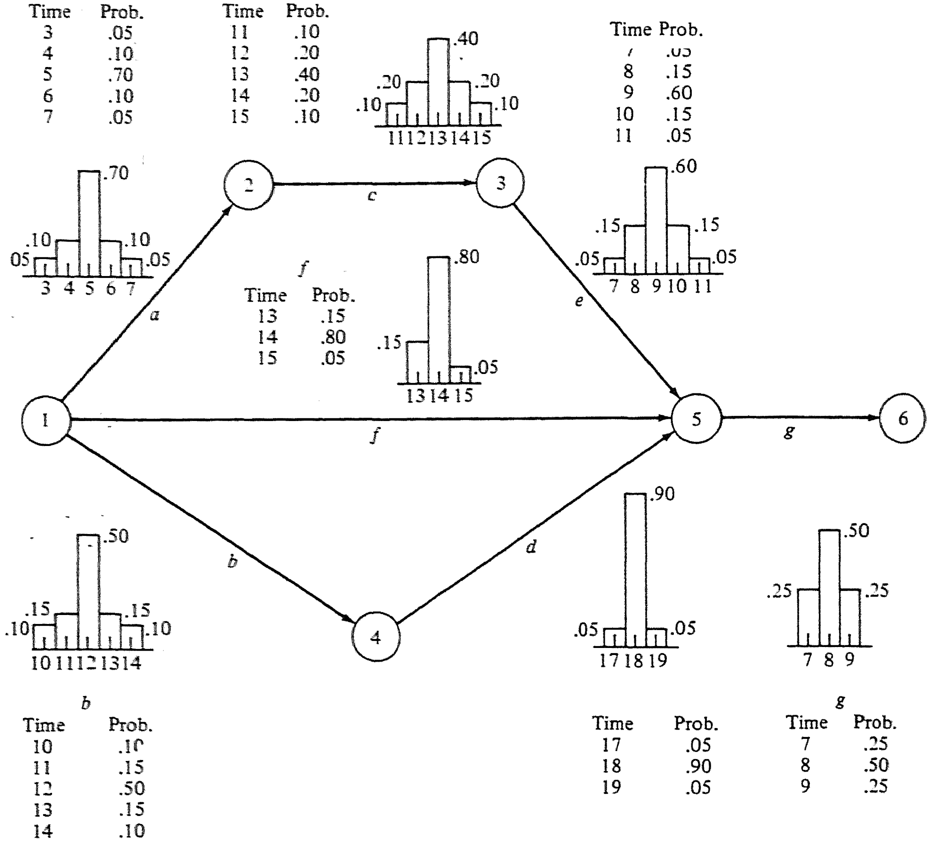


Fig. 7-10 Simulation of PERT network.

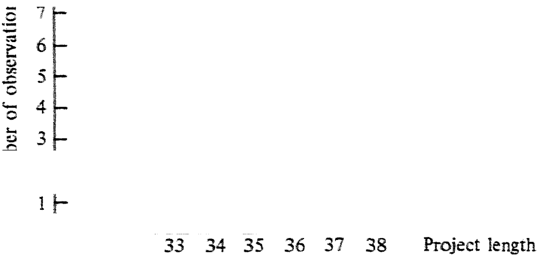


Fig. 7-11 Distribution of simulated completion dates.

Critical index Another useful piece of information that the simulation provides is a *critical index* for each activity. In Table 7-3 the critical path was recorded for each iteration. At the bottom of the table the percent of the time that each activity was critical is computed. This is called the "critical index" and is a measure of the likelihood that an activity will be critical. For example,

Table 7-3 Simulation results

	Critical path							
Simulation	*a*	*b*	*c*	*d*	*e*	*f*	*g*	*Completion date*
1	1		1		1		1	34
2		1		1			1	38
3		1		1			1	34
4		1		1			1	38
5	1	1	1	1	1		1	35
6	1	1	1	1	1		1	33
7		1		1			1	34
8	1	1	1	1	1		1	36
9	1		1		1		1	35
10	1		1		1		1	37
11		1		1			1	36
12		1		1			1	34
13		1		1			1	35
14	1		1		1		1	36
15	1		1		1		1	37
16	1		1		1		1	36
17	1	1	1	1	1		1	35
18	1		1		1		1	35
19	1		1		1		1	35
20	1		1		1		1	34
21		1		1			1	35
22	1		1		1		1	38
23	1	1	1	1	1		1	34
24	1	1	1	1	1		1	35
25	1	1	1	1	1		1	36
Critical index	.68	.60	.68	.60	.68	.00	1.00	

activity *a* has a critical index of .68. This means that if the project was undertaken repeatedly, approximately 68 percent of the time activity *a* would be found on the critical path. With these indices management knows which activities to watch closely, which to watch occasionally, and which are quite unlikely to become critical.

Quite noticeable, then, should be the fact that a critical path does not accompany a solution by simulation, but only a critical index for each activity. This, in fact, is more useful than the traditional approach to PERT which we first developed, for now we have taken account of critical and near-critical paths in our project plan.

CRITICAL PATH METHOD

The critical path method (CPM) was developed for situations where the project manager has the flexibility to channel additional resources to certain activities in the interest of reducing the project's completion date. It does not, however, have PERT's ability to incorporate uncertainty into the analysis. Therefore CPM is a deterministic model.

ACTIVITY AND PROJECT COSTS

The central idea in CPM is that certain costs vary with time. These costs are broken down into two categories: direct activity costs and indirect project costs. *Direct activity costs* are those costs which are incurred by the use of men, machines, and money. There is an inverse relationship between these costs and the length of time that it takes to complete an activity. In Fig. 7-12*a* it can be seen that as the expected completion date of an activity is pushed up, the direct cost of performing that activity increases. This follows since additional resources are necessary to accomplish a decrease in activity time.

The second set of costs is related to the *indirect costs* associated with the entire project. These include supervision, some supplies, and a host of fixed expenses which are normally allocated to the project. Many of these increase with the duration of the project. Therefore if the total project length can be shortened, these costs will be saved. The direction of these project-related costs is shown in Fig. 7-12*b*.

We have identified two sets of costs each working in opposite directions. If an activity is shortened, certain costs increase; whereas if the project is shortened, other costs decrease. It can therefore be concluded that additional resources can be applied to an activity provided that the cost of this additional effort is less than the savings which accrue from a shorter project cycle.

Consider the example shown in Fig. 7-13. The data above each arrow represent the average time or normal time for the completion of that activity, the minimum possible completion time if extra resources were employed, and finally, the cost of reducing the average time in "$/day." Activity *b*, for example, has a normal completion time of 4 days and can be completed in a minimum of 2 days if sufficient resources are used. The cost of reducing the completion time is $1 per day. Therefore if the activity time is reduced from 4 days to 3 days, an additional cost of $1 is incurred. If the time is reduced to 2 days, then the total cost of reduction is $2. It is not possible to reduce the time further.

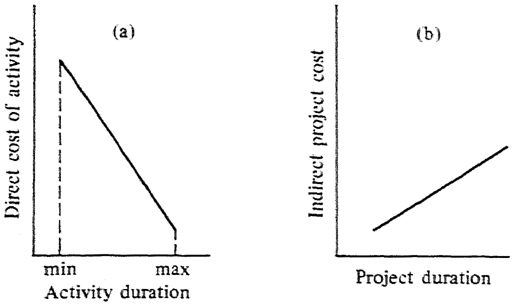


Fig. 7-12 Direct activity and indirect project costs.

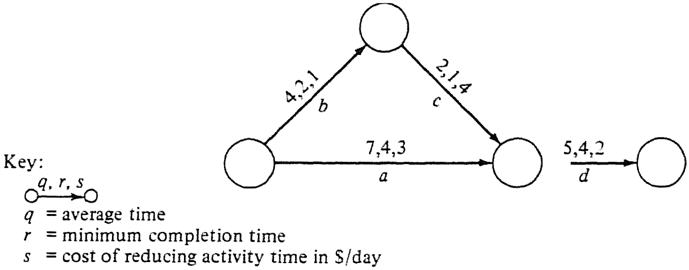

4,2,1
b
2,1,4
c
7,4,3
a
5,4,2
d
Key:
q, r, s
q = average time
r = minimum completion time
s = cost of reducing activity time in $/day

Fig. 7-13 Critical path method example.

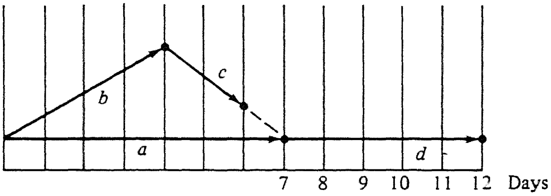


Fig. 7-14 Critical path method time graph.

CPM TIME GRAPH

The problem is solved on a CPM time graph. First a time scale is drawn, after which the activity arrows are entered at their earliest start time. This is illustrated in Fig. 7-14. The only way that this schedule can be reduced is by the commitment of additional resources. For each day that the schedule is reduced, a savings of $5 per day in indirect project costs is realized. Therefore if the project is shortened by 3 days, this indirect savings would total $15.

Now we can compare these direct costs and indirect project savings to see if any action should be taken. As it stands now, the project will be completed in 12 days. The critical path of the project is *a*-*d*, and to shorten the project length it only makes sense to focus on activities which lie on the critical path. Therefore, either activity *a* or activity *d* should be shortened. Activity *a* can be reduced to a minimum of 4 days at a cost of $3 per day, and *d* can be reduced to a minimum of 4 days at a cost of $2 per day. It is therefore more economical to reduce activity *d* to its minimum of 4 days. Since the project now takes 1 less day, the indirect cost saving is $5 and the net saving is $5 — $2 = $3.

If we ignore those costs which are not under our control at this stage in the planning process, the controllable cost of the schedule before any changes were made was:

$$\begin{aligned}\text{Cost} &= \frac{\text{Cost of reducing}}{\text{completion date}} + \text{indirect costs} \\ &= \qquad 0 \qquad + (12\text{ days})(\$5/\text{day}) \\ &= \$60\end{aligned}$$

Now that we have reduced the schedule to 11 days, the cost becomes

$$\text{Cost} = \$2 + 11(\$5) = \$57$$

We have therefore effected a \$3 cost savings. The new CPM graph reflecting this change is found in Fig. 7-15.

The critical path of our new CPM graph is still *a-d*; however, *d* can no longer be shortened. The cost of shortening *a* is \$3 per day and it *could* be shortened to 4 days; but after it is shortened by 1 day, two critical paths, *b-c-d* and *a-d*, occur. Moving one step at a time, we first reduce *a* by 1 day. The controllable cost of this new 10-day schedule now becomes

$$\begin{aligned}\text{Cost} &= (\$2 + \$3) + 10(\$5) = 5 + 50\\ &= \$55\end{aligned}$$

The first cost in the first parenthesis represents the \$2 cost that was incurred in the previous step, at which time we shortened the project from 12 to 11 days. The \$3 is the cost of shortening the project from 11 to 10 days. The new CPM graph is drawn in Fig. 7-16.

We now have two critical paths *a-d* and *b-c-d*. Since *d* cannot be shortened any further, any reduction in the project must now occur in activity *b* or *c* *and* in *a*. The choices we have, therefore, are to shorten *b* and *a* or *c* and *a*. The cost of shortening *b* and *a* is \$4 per day. The most that they can be shortened is by 2 days, since *b* cannot be completed in less than 2 days. If, on the other hand, we shorten *c* and *a*, the cost per day is \$7. We therefore choose to shorten *b* and *a*. The cost of the schedule is

$$\begin{aligned}\text{Cost} &= [\$2 + \$3 + 2(\$4)] + 8(\$5)\\ &= \$13 + \$40\\ &= \$53\end{aligned}$$

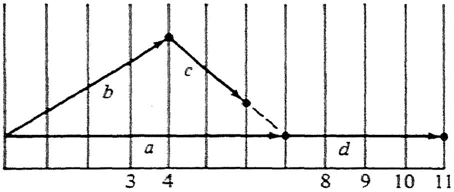


Fig. 7-15 Critical path method graph shortened to 11 days.

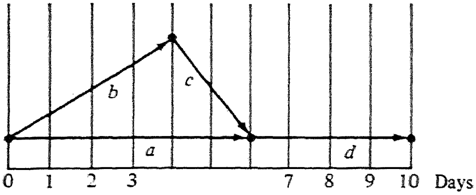


Fig. 7-16 Critical path method graph shortened to 10 days.

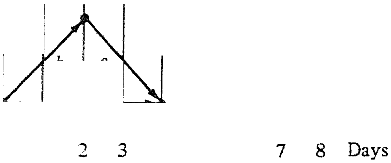


Fig. 7-17 Critical path method graph shortened to 8 days.

The new CPM graph is shown in Fig. 7-17. This is the schedule with the lowest cost, since any further reduction in project time must come as a result of decreasing activities *c* and *a* at a cost of $7 per day, which exceeds the savings that such a change would generate.

//// LINEAR PROGRAMMING FORMULATION OF A CPM NETWORK

In the PERT network the variable minimized was project time, and activity times were given. In the CPM network, however, the variable minimized was total cost, which was a function of both activity time t_{ij} and project length. Therefore we must consider activity time in the CPM linear programming model as a variable.

If we assume that the cost for reducing the activity time is linear, we can express it as

$$C_{ij} = a_{ij} - b_{ij}t_{ij}$$

where a_{ij} = intercept
$-b_{ij}$ = negative slope of cost function

As you recall, this function is defined only between the minimum activity time M_{ij} and the average time A_{ij}.

$$M_{ij} \leq t_{ij} \leq A_{ij}$$

In CPM there are two relevant costs, the direct activity costs and the indirect job costs. The objective is to minimize their sum. The total direct activity costs can be expressed as:

$$\sum_{\text{all } ij} (a_{ij} - b_{ij}t_{ij})$$

The indirect project costs are:

$$CX_m$$

where X_m = completion date of project
C = indirect costs per day

Therefore, if we minimize the sum of these costs, we have

$$\text{Minimize } C = \sum_{\text{all } ij} (a_{ij} - b_{ij}t_{ij}) + CX_m$$

Since all the a_{ij} are constant, we can rewrite this objective function in the following way:

$$\text{Minimize } C = -\sum_{\text{all } ij} b_{ij}t_{ij} + CX_m$$

The first set of constraints must specify that the difference between two connected events be at least as great as the duration of the activity.

$$X_j - X_i \geq t_{ij} \qquad \text{for all } ij$$

or

$$X_j - X_i - t_{ij} \geq 0 \qquad \text{for all } ij$$

The second set of constraints restricts the values of t_{ij} to a maximum of A_{ij}.

$$t_{ij} \leq A_{ij} \qquad \text{for all } ij$$

The third set of constraints ensures that the t_{ij} will be larger than M_{ij}.

$$t_{ij} \geq M_{ij} \qquad \text{for all } ij$$

Summarizing the model, we have

$$\text{Minimize } C = -\sum_{\text{all } ij} b_{ij}t_{ij} + CX_m$$

$$\begin{aligned} \text{Subject to} \qquad X_j - X_i - t_{ij} &\geq 0 \qquad \text{for all } ij \\ t_{ij} &\leq A_{ij} \qquad \text{for all } ij \\ t_{ij} &\geq M_{ij} \qquad \text{for all } ij \\ \text{All } X_{ij},\ t_{ij} &\geq 0 \end{aligned}$$

We will now proceed to formulate the problem illustrated in Fig. 7-13 as a linear programming model. For convenience the network is repeated in Fig. 7-18. First we formulate the objective function:

$$\text{Minimize } C = -1t_{12} - 3t_{13} - 4t_{23} - 2t_{34} + 5X_4$$

The first set of constraints can be formulated as:

$$\begin{aligned} X_2 - X_1 - t_{12} &\geq 0 \\ X_3 - X_1 - t_{13} &\geq 0 \\ X_3 - X_2 - t_{23} &\geq 0 \\ X_4 - X_3 - t_{34} &\geq 0 \end{aligned}$$

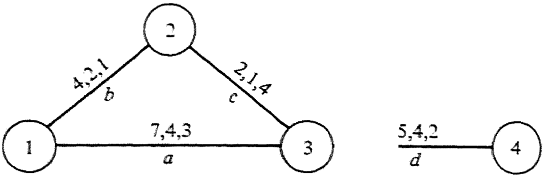


Fig. 7-18 Linear programming example.

Next the constraints that limit t_{ij} to a maximum of A_{ij}:

$$t_{12} \leq 4$$
$$t_{13} \leq 7$$
$$t_{23} \leq 2$$
$$t_{34} \leq 5$$

Finally the constraints that require values of t_{ij} at least as large as M_{ij}:

$$t_{12} \geq 2$$
$$t_{13} \geq 4$$
$$t_{23} \geq 1$$
$$t_{34} \geq 4$$

Solving this problem by a standard linear programming code we get

$$t_{12} = 2 \qquad X_1 = 0$$
$$t_{13} = 4 \qquad X_2 = 2$$
$$t_{23} = 2 \qquad X_3 = 4$$
$$t_{34} = 4 \qquad X_4 = 8$$

This agrees with our previous solution. ////

SCHEDULING WITH LIMITED RESOURCES

The analysis up to this point has been based on the assumption that if its predecessors were completed, an activity could begin. Underlying this assumption is the fact that sufficient resources will be available to start each activity at its earliest start date. Consider the network shown in Fig. 7-19. It is drawn on a time graph from which it can be seen that the critical path is *c-f-g* and that the earliest this project can be completed is 10 days.

The figures above the activities represent a dimension to the problem which has not yet been considered: the resource requirement of that activity. A resource could be any number of things, including work force size, number of pieces of equipment, number of machines, and money. In this example the work force size will be that resource which is considered to be in scarce sup-

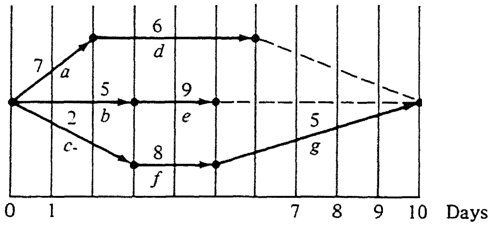


Fig. 7-19 Scheduling with limited resources.

ply. Therefore the figures above each activity represent the work force size that is required to undertake that activity.

If the project is executed according to the schedule of Fig. 7-19, the peak work load demand of 23 men occurs on the fourth and fifth days. As you can see from Fig. 7-20, the demand on other days drops off significantly. It may be that on any one day all 23 men are not available. Even if they were, it would seem to be more efficient if the demand for manpower could be smoothed by taking advantage of slack time that exists in noncritical paths. This would avoid costs of hiring and firing, overtime, and idle capacity.

HEURISTIC APPROACH TO RESOURCE SMOOTHING

In a simple case such as this, where but one resource must be considered, a simple heuristic procedure can be used to smooth the demand for the resource. First we focus on the day of peak demand and adjust an activity which has some slack. If we delay activity *e* by 3 days, the peak demand is reduced to a work force size of 14. The new network and load chart are shown in Fig. 7-21.

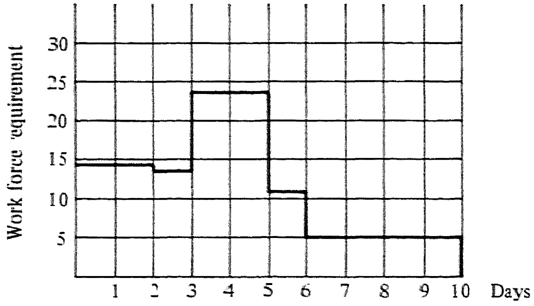


Fig. 7-20 Distribution of work load.

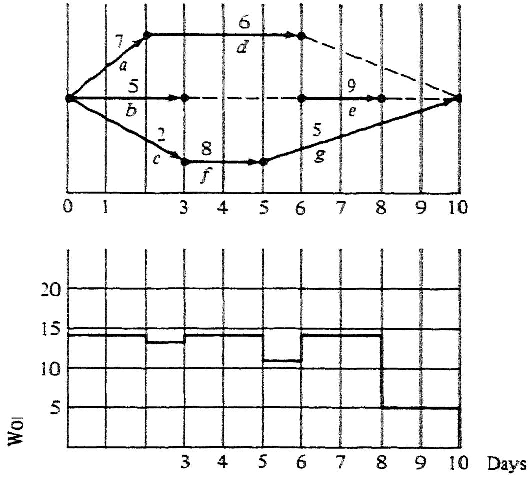


Fig. 7-21 Improved solution.

This rather informal approach to solving the problem becomes more difficult as we not only add complexity to the network but also introduce additional kinds of resources which are in scarce supply. If you do not believe this, try drawing a network with 15 or so activities where three resources are required by each activity. After you have drawn several networks and have been driven mad as the reduction in one resource leads to the violation of others, you will have an intuitive feel for the extent of the problem.

//// AN INTEGER PROGRAMMING FORMULATION OF PROJECT SCHEDULING WITH LIMITED RESOURCES

Now we turn to one way in which an optimal solution to this problem can be achieved. Here we will present a model developed by Wiest and Levy.[1] In the model we will utilize a decision variable X_{jd} which has the following interpretation:

$$X_{jd} = \begin{cases} 1 & \text{if activity } j \text{ is performed on day } d \\ 0 & \text{otherwise} \end{cases}$$

Therefore the first set of constraints must include

$$0 \leq X_{jd} \leq 1 \qquad \begin{array}{l} j = 1, \ldots, n \\ d = 1, \ldots, z \\ \text{all } X_{jd} \text{ limited to integers} \end{array}$$

In addition to being constrained between 0 and 1, the further restriction that all X_{jd} are limited to integers allows only 0,1 values and requires that we use integer programming codes.

The second set of constraints ensures that all activities will be performed and completed.

$$\sum_{d=1}^{z} X_{jd} = t_j \qquad j = 1, \ldots, n$$

where t_j = number of days required to complete activity j

According to this constraint, activity j must be scheduled for exactly t_j days.

The third set of constraints specifies that the capacity of each scarce resource will not be exceeded each day.

$$\sum_{j=1}^{n} C_{rj}X_{jd} \leq A_{rd} \qquad \begin{array}{l} d = 1, \ldots, z \\ r = 1, \ldots, m \end{array}$$

where C_{rj} = quantity of resource r required in performing activity j
A_{rd} = maximum quantity of resource r available on day d

---

[1] Jerome Wiest and Ferdinand Levy, *A Management Guide to PERT/CPM*, Prentice-Hall, Inc., Englewood Cliffs, N.J., 1969, p. 124.

The fourth set of constraints ensures that all predecessors must be completed before an activity is begun.

$$t_p X_{jd} \le \sum_{i=}^{d-1} X_{pi} \qquad \begin{array}{l} \text{all } p \in P_j \\ d = 1, \ldots, z \\ j = 1, \ldots, n \end{array}$$

First we must define p. Let p represent an activity which is an immediate predecessor to j. All immediate predecessors to activity j will be said to be contained in set P. Therefore, we can write $p \in P_j$, which can be interpreted as all the immediate predecessors of activity j. Since this set of constraints may be difficult to understand, an example should help.

Let us turn our attention to one subset of these precedence constraints. We will analyze the constraints that must be developed for the fifth activity on days 1, 2, and 3. Furthermore, we assume that activities 1 and 3 are the immediate predecessors of activity 5. Constructing these constraints, we have

$$\text{For} \begin{cases} p = 1 \\ d = 1 \\ j = 5 \end{cases} \quad t_1 X_{51} \le 0 \qquad \text{For} \begin{cases} p = 3 \\ d = 1 \\ j = 5 \end{cases} \quad t_3 X_{51} \le 0$$

$$\text{For} \begin{cases} p = 1 \\ d = 2 \\ j = 5 \end{cases} \quad t_1 X_{52} \le X_{11} \qquad \text{For} \begin{cases} p = 3 \\ d = 2 \\ j = 5 \end{cases} \quad t_3 X_{52} \le X_{31}$$

$$\text{For} \begin{cases} p = 1 \\ d = 3 \\ j = 5 \end{cases} \quad t_1 X_{53} \le X_{11} + X_{12} \qquad \text{For} \begin{cases} p = 3 \\ d = 3 \\ j = 5 \end{cases} \quad t_3 X_{53} \le X_{31} + X_{32}$$

The two constraints across the top row require that activity 5 can begin on day 1, $X_{51} = 1$, only if its immediate predecessors, activities 1 and 3, require zero time, $t_1 = 0$ and $t_3 = 0$. It is only under this condition that the constraints will not be violated. For example, if activity 1 took 1 day, $t_1 = 1$, and activity 5 was performed on day 1, $X_{51} = 1$, the constraint would be violated, for $1 \nless 0$.

The last row of constraints can be interpreted in a similar fashion. If activity 5 is started on day 3, $X_{53} = 1$, its immediate predecessors, activities 1 and 3, must be fully completed by day 2. Suppose predecessor 1, $p = 1$, requires $t_1 = 2$ days, and activity 5 is scheduled to begin on the third day, $X_{53} = 1$. If activity 1 has been scheduled on day 1, $X_{11} = 1$, and day 2, $X_{12} = 1$, the plan thus far is feasible, since

$$t_1 X_{53} \le X_{11} + X_{12}$$

$$2(1) \le 1 + 1$$

Since precedence constraints must be written for all immediate predecessors of each activity, and for each day, you can see where the size of the problem can get quickly out of hand.

Next we turn to the last set of constraints, which ensures that no jobs will be split.

$$t_jX_{jd} - t_jX_{j(d+1)} + \sum_{i=d+2} X_{ji} \le t_j \qquad \begin{aligned} d &= 1, \ldots, z \\ j &= 1, \ldots, n \end{aligned}$$

A *split* is defined as performing part of the activity on day 2 and the rest of the activity on days 7 and 8. Again we will use an example to illustrate this constraint set. Consider an activity $j = 6$ which requires $t_6 = 3$ days for completion. The maximum number of days that the entire project could take is estimated at $z = 8$ days. If the activity is split between days 2, 7, and 8, the constraint associated with day z can be written as

$$\begin{aligned} t_6X_{62} - t_6X_{63} + (X_{64} + X_{65} + X_{66} + X_{67} + X_{68}) &\le t_6 \\ 3(1) - 3(0) + (0 + 0 + 0 + 1 + 1) &\le 3 \\ 5 &\nleq 1 \end{aligned}$$

It is clear that this split has violated the constraint. The only way in which the set of constraints associated with each activity will not be violated is if the activities are not split. The task of constructing this set of constraints is also sizable, since one must be constructed for each activity and for each day. The objective function is written as follows:

$$\text{Minimize } 1 \sum_{j=1}^{n} X_{jk} + 4 \sum_{j=1}^{n} X_{j(k+1)} + 16 \sum_{j=1}^{n} X_{j(k+2)} + \cdots + R_z \sum_{j=1}^{n} X_{jz}$$

where $0 < k < z$

$$R_z = 4R_{(z-1)}$$

The R's serve as a cost explosion making the assignment of an activity to a later day more and more costly. This, of course, forces the completion date to be as early as possible. ////

SUMMARY

Each network model that has been presented has its own set of assumptions. The choice of a model, in any one situation, depends upon the degree to which that model incorporates the predominant characteristics of the real-world problem. If, for example, there are limited resources, the PERT and CPM models would be inadequate by themselves. If the project is plagued with uncertainty, a PERT model would seem appropriate.

If resources are available to shorten the project time, CPM methods can be used to minimize the total cost of the project. The assumption, however, in the CPM model that these costs are linear may prove to be too severe.

The user, then, has a difficult time in selecting the appropriate model, since it is unlikely that any one will exactly meet his requirements. Therefore, when the outcome of a model is in turn used to solve the real-world problem, some discretion may have to be exercised.

QUESTIONS

Q7-1. What is the difference between PERT and CPM?

Q7-2. What is a dummy activity and when must it be employed?

Q7-3. What is the significance of a critical path?

Q7-4. How is slack time computed?

Q7-5. As the result of a PERT analysis, the expected completion time of a project is computed to be 60 days with a standard deviation of 4 days. What can be said about the likelihood of an actual completion time in excess of 70 days?

Q7-6. Suppose that the outcome of a PERT analysis is such that the expected completion time for the critical path is 45 days with a standard deviation of 3 days. There is, however, another path whose expected completion time is 42 days with a standard deviation of 4 days. What would be the problem if the critical path was used to make probability statements about project completion?

Q7-7. Why might it be necessary to employ simulation in PERT analysis?

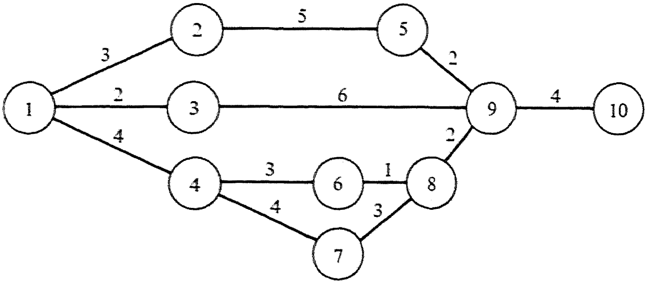
Q7-8. Is the outcome of a conventional PERT analysis any different from a simulated PERT analysis? Explain.

Q7-9. In CPM, trade-offs can be made between two kinds of costs. Explain.

Q7-10. Why is it more difficult to solve a critical path problem with scarce resource

PROBLEMS

P7-1. Find the critical path and slack times for the following network.



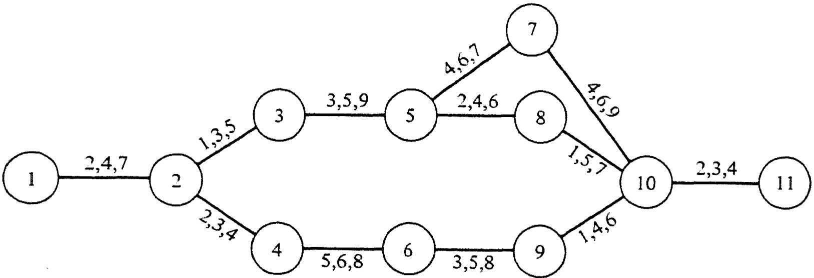
P7-2. Draw the network and find the critical path and slack times for the following problem.

Activity	*Immediate predecessor*	*Time*
1,2	—	5
1,3	—	7
2,4	1,2	4
3,4	1,3	3
4,7	2,4	6
3,5	1,3	4
5,10	3,5	1
7,9	4,7	6
4,6	2,4 and 3,4	7
6,8	4,6	2
7,8	4,7	4
8,10	6,8 and 7,8	5
9,10	7,9	6
10,11	9,10 and 8,10 and 5,10	2

P7-3. Draw the network for the following project and find the critical path.

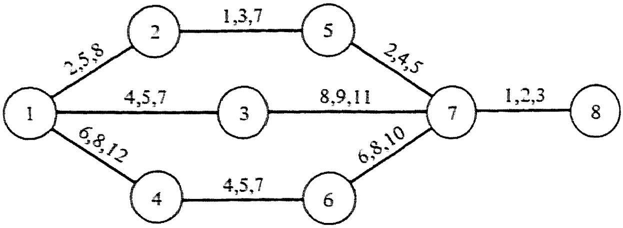
Activity	*Immediate predecessor*	*Time*
1,2	—	5
1,3	—	4
2,4	1,2	6
3,4	1,3	2
4,5	2,4	1
4,6	2,4 and 3,4	7
5,7	4,5	8
6,7	4,6	4
7,8	6,7 and 5,7	3

P7-4. Find the critical path of the following PERT network.

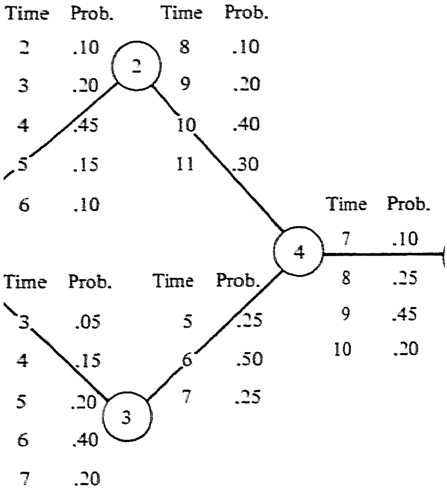


P7-5. What is the likelihood that the critical path found in Prob. 7-4 will take longer than 30 days?

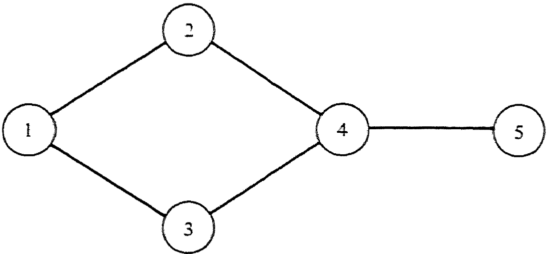
P7-6. Find the critical path of the following PERT network and determine the likelihood that this path will take more than 21 days.



P7-7. Simulate the following PERT network. Base your distribution of completion dates on 20 replications. What is the likelihood that each activity will become critical?



P7-8. The design of a new production system requires the completion of five major activities. This network is shown below. The supervisor of the project has subjectively estimated the possible completion times for each of these activities. On the basis of 20 replications, estimate the likelihood that this project will take more than 12 weeks. What is the likelihood that activity 1-2 will be critical?



Activity 1-2

Time	*Probability*
3	.12
4	.35
5	.23
6	.17
7	.13

Activity 1-3

Time	*Probability*
1	.10
2	.20
3	.45
4	.25

Activity 2-4

Time	*Probability*
4	.10
5	.20
6	.60
7	.10

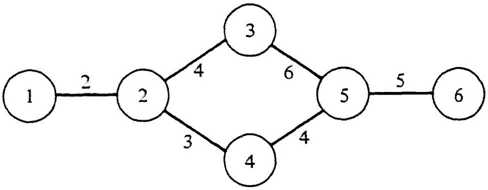
Activity 3-4

Time	*Probability*
5	.10
6	.20
7	.35
8	.25
9	.10

Activity 4-5

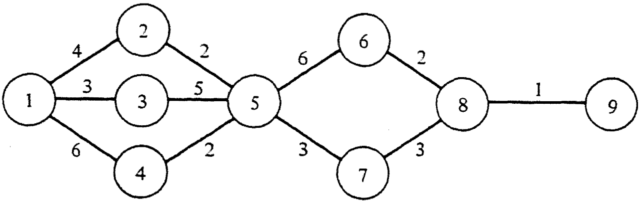
Time	*Probability*
1	.20
2	.25
3	.35
4	.20

P7-9. Formulate the following as a linear programming problem. Set up only.



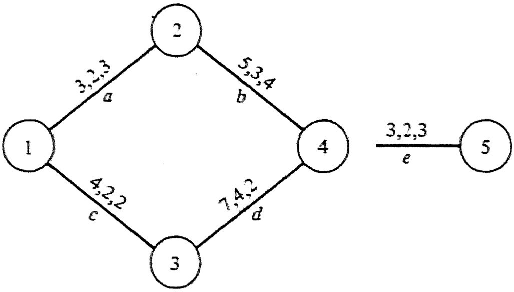
P7-10. Solve Prob. 7-9 using a standard LP computer code.

P7-11. Formulate the following as a linear programming problem. Set up only.



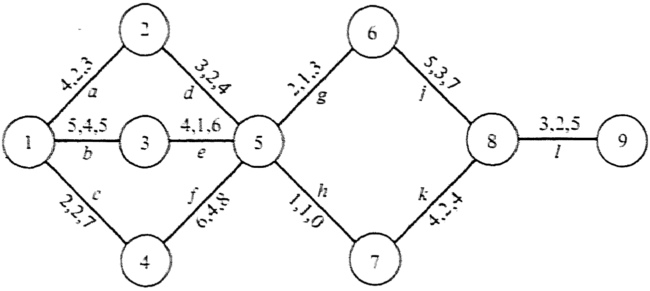
P7-12. Solve the following CPM problem.

The figures above each activity represent the average time in which the activity can be completed, the minimum time, and the cost of reducing activity time in dollars per day. In addition, the savings which accrue if the project is shortened is $5 per day.



P7-13. Solve the following CPM problem.

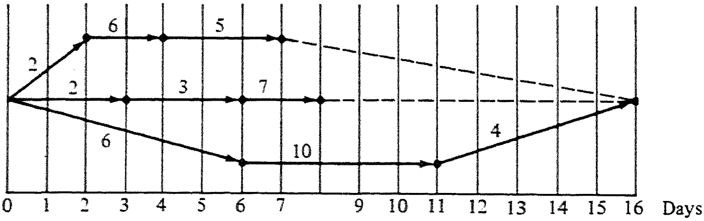
For each day that the project is shortened, a savings of $7 is realized.



P7-14. Formulate Prob. 7-12 as a linear programming problem.

P7-15. Formulate Prob. 7-13 as a linear programming problem.

P7-16. The major resource which a firm requires to complete the following network is labor. Since they would like to smooth the demand for this resource, determine a plan whereby this is accomplished as well as possible.



Figures above each activity indicate the quantity of the scarce resource required.

BIBLIOGRAPHY

Davis, Edward W.: "Resource Allocation in Project Network Models—A Survey," *Journal of Industrial Engineering,* April 1966.

——— and George E. Heidorn: "An Algorithm for Optimal Project Scheduling under Multiple Resource Constraints," *Management Science,* vol. 17, no. 12, August 1971, pp. B803–B816.

Levy, F. K., G. L. Thompson, and J. D. Weist: "The ABC's of the Critical Path Method," *Harvard Business Review,* vol. 41, no. 5, September 1963, pp. 98–108.

Miller, R. W.: "How to Plan and Control with PERT," *Harvard Business Review,* vol. 40, no. 2, March–April 1962, pp. 93–104.

Odom, Ralph G., and E. Blystone: "A Case Study of CPM in a Manufacturing Situation," *Journal of Industrial Engineering,* vol. 15, no. 6, November–December 1964.

Schoderbek, P. P.: "A Study of the Applications of PERT," *Journal of Academy of Management,* vol. 8, no. 3, September 1965, pp. 199–210.

Siemens, Nicolai: "A Simple CPM Time-Cost Tradeoff Algorithm," *Management Science,* vol. 17, no. 6, February 1971, p. B354.

Wiest, Jerome D., and Ferdinand K. Levy: *A Management Guide to PERT/CPM,* Prentice-Hall, Inc., Englewood Cliffs, N.J., 1969.

8

Facilities Layout

INTRODUCTION

Facilities layout is closely tied to the macro, micro, and behavioral design of the process. For it is only after the process has been designed, the make or buy decision made, the sequence of activities determined, the equipment selected, and the materials handling problem considered that plant layout usually begins. In other words, plant layout is really a continuation of process design, and as such it should be subjected to the same iterative process. As the attention shifts from macro to micro to behavioral and then to plant layout, decisions made at these earlier stages constrain the way in which the plant can be laid out. If an existing plant will be used, then its physical dimensions and facilities will impose additional constraints on the layout. Therefore, what seemed like a feasible design as it progressed through the macro, micro, and behavioral stages might prove infeasible when the layout of the existing plant is considered. The designers must then go back and make the necessary changes in these earlier stages, or perhaps even consider changes in the physical plant. It should be made clear that although we focus on each of these stages separately in the text, they are usually considered sequentially in the design process, although sometimes a select few are considered simultaneously.

PRODUCT AND PROCESS LAYOUTS

PRODUCT LAYOUT

One way in which to categorize a layout is by product or by process. A product layout takes the form of a production line and can be justified only when planned output is high and the product specifications are certain. The stations and equipment are devoted exclusively to a single product; and since they are usually special-purpose, the fixed costs are high and the process is inflexible. Therefore a high volume must be anticipated to defray this cost; and because of the inflexibility of the equipment, the product specifications should not be expected to change by very much over the life of the equipment.

A consequence of these characteristics is that the layout problem for product-oriented lines is typically small. Once the process decision has been made, there is usually little area for discretion in the location of the equipment; layout, for the most part, is determined by technology itself.

For example, a large electroplating department is considering the installation of an automatic chrome-plating machine with 25 different stations. Each station contains either a chemical cleansing solution, an electroplating solution, or a water rinse. The part to be chrome-plated is mounted on an overhead conveyor which moves both horizontally and vertically. It drops the part into the solution, raises it when the plating time is up, moves it to the next station, lowers it, and so on. If the machine is purchased, it is clear that there will not be a sizable layout problem. The manufacturer of the plating equipment offers but a few alternative layouts for the process. The stations can either be strung out in a straight line or arranged in the form of the letter U. Beyond this basic choice there is little more that needs to be done in laying out this facility.

Product layouts need not be highly capital-intensive. They can embody substantial levels of labor. Many times, for example, the final assembly of a product is laid out as a labor-intensive assembly line. Workers, therefore, perform the necessary operations at each station with the help of tools and fixtures. To determine if such a strategy is efficient, the line must first be balanced, perhaps using the techniques presented in Chapter 6. Then the cost consequence of this strategy can be compared with the cost of others. If the product layout is chosen, the outcome of the line-balancing analysis will play a dominant role in the physical layout; again there will be little room for discretion.

We can therefore conclude that the facilities layout problem associated with product layouts is not large. In fact, graphic aids such as two- and three-dimensional drawings, as well as two- and three-dimensional models, can be used to scrutinize the possible alternatives with little risk of being too superficial. This is not the case for process layouts. They are more complex to analyze and design. Consequently the rest of this chapter is devoted to process layout problems.

PROCESS LAYOUT

Characteristic of process layouts is the fact that similar work stations are grouped together. In a job shop, for example, we find a work center containing a group of lathes, another center with a group of milling machines, and so on. Other examples include a hospital laboratory, a school's language lab, and an office typing pool.

In a process layout many different jobs are processed through each machine. One job by itself could not justify the exclusive use of the stations through which it is routed. Therefore a product layout would be uneconomical, and a process layout in which the job shares facilities with others is a much more effective use of scarce resources.

With the job moving from one work center to the next, material handling costs are higher than they would be under a product layout strategy. As a result, a major consideration in process layout is the design of a layout which minimizes transportation costs for the jobs which are processed through the facility.

The structure of this chapter will be to first analyze the problem of placing a facility in a location when the consequence of such a placement is totally independent of the facilities which occupy other locations. We call this problem *process layout:independence.* To solve such a problem, a branch and bound assignment algorithm is developed. Then the attention will shift to facility layout problems where the consequence of placing a facility in a location is dependent on the location of other facilities. Two methods for solving this problem are presented. The first is a mathematical programming model whose size turns out to be quite substantial. To save the day, a heuristic is introduced which finds very good but not optimal solutions to the problem. Those who wish to avoid the mathematical programming model may do so with only a modest loss to the continuity of the chapter.

PROCESS LAYOUT: INDEPENDENCE

Sometimes the consequence of placing a facility in a location is totally independent of the facilities which occupy other locations in the immediate vicinity. As an example of this, consider the problem faced by the Ace Steel Company. They produce five steel products, including steel cable, I beams, sheet steel, bar stock, and steel rods used in reinforced concrete. To store these products after they have been manufactured, they have set aside an area which has been divided into five locations. The problem that must be solved is this: Which product or "facility" do we assign to each location?

The choice of assignment depends upon the criterion chosen. Here it seems logical to choose as that criterion the minimization of material handling cost.

When one of these five items is demanded, a forklift truck is dispatched to the appropriate location. It then moves the item to the shipping ramp or onto a waiting delivery truck. The handling costs associated with this activity therefore depend only upon the zone to which that product is assigned.

Location

Facility	1	2	3	4	5	
A	5	12	15	13	17	1
B	8	10	6	11	7	1
C	9	3	11	8	9	1
D	2	5	9	4	5	1
E	5	4	4	3	13	1
	1	1	1	1	1	

Fig. 8-1 Independent facilities layout problem.

These costs for all possible combinations of facilities and locations have been determined and are recorded in Fig. 8-1. For example, if facility A is assigned to location 1, the yearly cost of handling will be \$5(000). If facility A is, instead, assigned to location 2, the cost will be \$12(000), and so on.

In addition to these costs, the constraints associated with this problem are also recorded in Fig. 8-1. The 1s located below the last row imply that each location should have assigned to it one and only one facility. The 1s located to the right of the last column imply that each facility should be assigned to one and only one location. These constraints are called *rim conditions*. We can now say that the problem is to assign facilities to locations in such a way as to minimize material handling costs while observing the rim conditions.

Let us look at one alternative assignment.

Facility	A	B	C	D	E
Location					

Assignment	*Cost*
A-3	15
B-2	10
C-1	9
D-4	4
E-5	13
	51

The cost of this alternative is 51 (that is, \$51,000). There are, however, many other alternatives. In fact, there are 5!, or 120, of them. To compute the cost for each and then from this set select the lowest-cost alternative would be inefficient. There is, indeed, a better way.

A MATHEMATICAL-PROGRAMMING FORMULATION OF THE PROBLEM

The general form of this problem can be categorized as a linear assignment problem. In its formal description, a decision variable X_{ij} is defined which can take on the following values:

$$X_{ij} = \begin{array}{ll} 1 & \text{if facility } i \text{ is assigned to location } j \\ 0 & \text{otherwise} \end{array}$$

The objective function can be written as:

$$\text{Minimize } C = \sum_{i=1}^{n} \sum_{j=1}^{n} C_{ij}X_{ij}$$

where C_{ij} = cost of assigning facility i to location j

Another requirement of this model is that the number of facilities n must be equal to the number of locations n. This is expressed by the limits above the two summation signs.

The constraints include a set which ensures that each facility is assigned to one and only one location.

$$\sum_{j=1}^{n} X_{ij} = 1 \qquad i = 1, 2, \ldots, n$$

The second set ensures that each location is assigned one and only one facility.

$$\sum_{i=1} X_{ij} = 1 \qquad j = 1, 2, . .$$

Location

			3	4	5
A	5	12	15	13	17
B	8	10	6	11	7
C	9	3	11	8	9
D	2	5	9	4	5
E	5	4	4	3	13

Fig. 8-2 The amount by which each column can be reduced.

Location

		3	4	
3	9	11	10	12
6	7	2	8	2
7	0	7	5	4
0	2	5	1	0
3	1	0	0	8

Fig. 8-3 After column reduction.

In addition we must add:

$$X_{ij} = \text{integer value} \qquad i = 1, 2, \ldots, n$$
$$j = 1, 2, \ldots, n$$

The problem illustrated in Fig. 8-1 would be expressed in the following way.

$$\text{Minimize } C = 5X_{A1} + 12X_{A2} + 15X_{A3} + 13X_{A4} + 17X_{A5} + 8X_{B1} + \cdots + 13X_{E5}$$

$$X_{A1} + X_{A2} + X_{A3} + X_{A4} + X_{A5} = 1$$
$$X_{B1} + X_{B2} + X_{B3} + X_{B4} + X_{B5} = 1$$

$$X_{E1} + X_{E2} + X_{E3} + X_{E4} + X_{E5} = 1$$
$$X_{A1} + X_{B1} + X_{C1} + X_{D1} + X_{E1} = 1$$
$$X_{A2} + X_{B2} + X_{C2} + X_{D2} + X_{E2} = 1$$

$$X_{A5} + X_{B5} + X_{C5} + X_{D5} + X_{E5} = 1$$
$$X_{ij} = \text{integer value} \qquad i = \text{A, B, C, D, E}$$
$$j = 1, 2, 3, 4, 5$$

It is quite possible to solve this problem using standard integer programming codes; however, it is much more efficient to use the linear assignment algorithm, to which we will shortly turn our attention.

SOLUTION BY OBSERVATION

To solve this problem an opportunity cost matrix is developed. This is accomplished by subtracting the lowest cost in each column from the other costs in that column. For example, in column 1 of Fig. 8-2 the lowest cost is 2. After reducing all the costs in that column by 2, we have the new column 1 of Fig. 8-3. Column reductions continue for the other four columns.

Interpreting the opportunity cost matrix of Fig. 8-3, we can see that if facility D is assigned to location 1, we will experience a zero opportunity cost. In other words, for location 1 the best possible assignment is facility D. If, however, we assign facility A to location 1, the cost is $3 *higher* than its best assignment. Clearly, if the zeros were positioned in such a way that each location could be assigned one of the five different facilities utilizing only zero cells, we would have the best solution. Can we accomplish this in our problem? Location 1 can be assigned facility D, location 2 can be assigned facility C, location 3 assigned facility E, but location 4 cannot be assigned

Location

ility

			3	4	5
A	0	6	8	7	9
B	4	5	0	6	0
C	7	0	7	5	4
	0	2	5	1	0
	3	1	0	0	8

Fig. 8-4 Opportunity cost matrix.

facility E, since it has already been assigned to location 3. Also, location 5 cannot be assigned facility D, since that has been assigned to location 1. Therefore, we have not solved the problem—yet.

Next we will subtract the lowest cost from each row and thereby generate an opportunity cost matrix with respect to facilities and locations. Therefore, from our last matrix, Fig. 8-3, we subtract the lowest cost in each row from the other costs in the row. The new matrix is presented in Fig. 8-4. We now have a matrix which has been reduced twice, once by columns and once by rows. There is therefore at least one zero in every column and every row. Once again we check to see if there exists a feasible assignment of facilities to locations in zero cells. There is, and it includes A-1, B-3, C-2, D-5, and E-4. The cost of this assignment is $5 + 6 + 3 + 5 + 3 = \$22(000)$. There is no other feasible assignment with a lower cost.

This example was a relatively simple one where the solution was identified after row and column reductions. In most cases, however, the location of the zeros after row and column reduction is *not* such that a solution can be identified. In those cases we must rely on the *assignment algorithm* for a solution technique.

SOLUTION BY ASSIGNMENT ALGORITHM

Consider the example shown in Fig. 8-5. The result of column reductions is shown in Fig. 8-6 and row reductions in Fig. 8-7. It can be seen that there is no feasible assignment of facilities to locations, using only zero cells. Other nonzero cells will have to be used. To establish which of these to use, we must employ a branch and bound assignment algorithm.

Branching technique The solution method which will be followed takes a problem and replaces it with two less complex problems. It then takes these new problems and replaces each of them with two more problems which are even less complex, and so on, until a final optimal solution is reached. To accomplish this, we start with the original problem, which implicitly contains

Location

Facility			3	4	5
A	9	11	9	12	10
B	22	26	23	24	25
C	13	15	13	16	14
D	5	8	11	13	9
E	15	17	14	16	17
				12	

Fig. 8-5 Independent facilities layout problem.

Location

	1	2	3	4	5	
	4	3	0	0	1	
B	17	18	14	12	16	12
C	8	7	4	4	5	
D	0	0	2	1	0	
	10	9	5	4	8	

Fig. 8-6 After column reduction.

Location

Facility	1	2	3	4	
A	4	3	0	0	1
B	5	6	2	0	4
C	4	3	0	0	1
D	0	0	2	1	0
	6	5	1	0	4

Fig. 8-7 After column and row reductions.

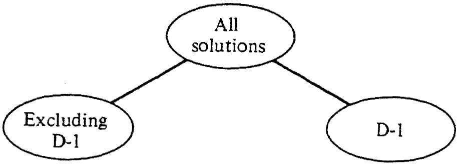


Fig. 8-8 Branching technique illustrated.

all possible solutions. Then we make an assignment of a facility to a location. It serves as the basis for partitioning the problem into these less complex problems. To illustrate this process, suppose that we decide that the best first assignment in our example is D-1. At that point we can reduce the problem to two smaller problems, one which contains all solutions which include D-1 as an assignment and another which explicitly eliminates D-1 from all possible solutions. This branching concept is illustrated in Fig. 8-8. The method then continues by branching from these reduced problems until a solution is reached.

Bounds Fortunately it will not be necessary to follow each branch in this tree. Only those branches which can clearly lead to the lowest-cost solution need be followed. To identify which these are, it is necessary to compute the lower bound for each branch. The lower bound represents the lowest possible cost of a solution *if* that branch is followed. It can be computed by summing all column and row reductions that were necessary up to and including that branch.

To begin, we must establish a lower bound for all possible solutions. This is accomplished by reducing the original problem by columns and rows.[1] Then the sum of these reductions is computed. In our example, column and row reductions were taken in Figs. 8-5 and 8-6. Their sum is 43 + 20 = 63. It is entered at the top of Fig. 8-9.† From this we can establish that there is no possible solution to the problem that will have a cost lower than 63. In fact, the only time that the optimal solution would be as low as 63 is when the zero cells are located in such a way that a feasible assignment can be made to them. We have already established that this is not the case for our problem.

---

[1] If you prefer, you can reduce first by rows and then by columns in this and all subsequent steps. The final answers will have the same cost.

† It will prove very helpful for studying the rest of this section if you reproduce this figure for reference purposes.

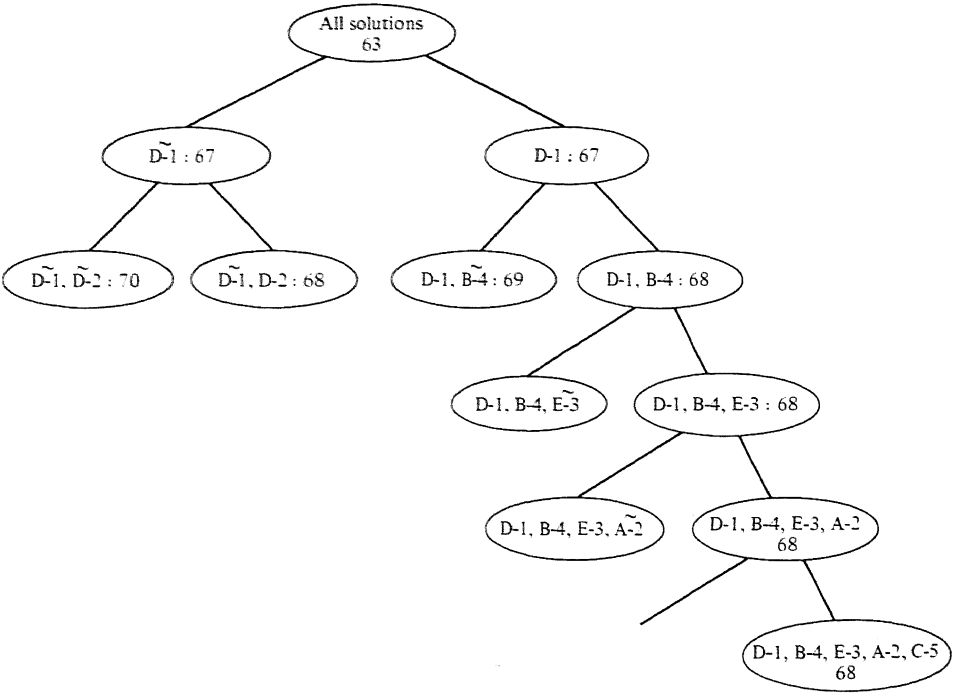
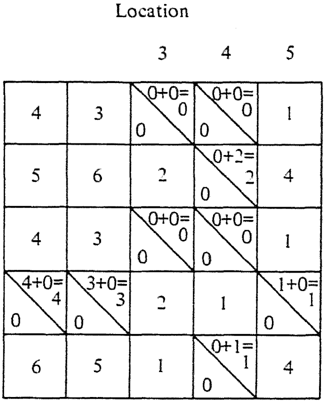


Fig. 8-9 Branching to a solution.



Location

		3	4	5
4	3	0+0= 0 0	0+0= 0 0	1
5	6	2	0+2= 2 0	4
4	3	0+0= 0 0	0+0= 0 0	1
4+0= 4 0	3+0= 3 0	2	1	1+0= 1 0
6	5	1	0+1= 1 0	4

Fig. 8-10

Now we turn to making our first assignment. It will serve as the basis for the first two branches.

Opportunity costs We make the first assignment to the zero cell which has the highest opportunity cost of *not* making an assignment to it. To accomplish this we must compute such a cost for each zero cell. So that our bookkeeping will be as organized as possible, the fully reduced matrix of Fig. 8-7 is redrawn with slash marks dividing each zero cell. This is done in Fig. 8-10. Below the slash mark the zero is entered, and above it we will place the opportunity cost of not making an assignment to that cell.

First we will focus upon D-1. If location 1 does *not* have facility D assigned to it, it must be assigned some other facility. In the interest of cost minimization, either one of facilities A and C could be assigned to location 1 at an additional cost of 4. Now, if D is no longer assigned to location 1, it must be assigned to some other location. Again in the interest of cost minimization, the lowest-cost alternative is to assign it to either of locations 2 and 5 at an incremental cost of zero. We can therefore conclude that the opportunity cost of *not* making an assignment of a facility to a location is the *sum* of the incremental cost of making another assignment to that location and that of assigning that facility to another location. In this case the opportunity cost is $4 + 0 = 4$. It is entered above the slash mark in that cell.

Next we evaluate the opportunity cost of not making an assignment to E-4. The lowest incremental cost of making another assignment to location 4 is zero, since any of facilities A, B, and C could be assigned to it. The lowest incremental cost of assigning E to another location is 1, since E could be assigned to location 3. Therefore the opportunity cost of not making the assignment E-4 is $0 + 1 = 1$.

Each zero cell must be evaluated separately, and the opportunity cost entered above the slash mark. The remaining opportunity costs are entered in Fig. 8-10.

The first assignment The assignment should be made to that cell with the highest opportunity cost. In this way we avoid incurring this high cost if an assignment is not made. The first assignment is therefore made in cell D-1.

The first two branches The first assignment serves as the basis for partitioning the problem into two less complex problems. This partitioning process is shown at the top of Fig. 8-9. One branch contains those solutions which include D-1 as an assignment and the other branch includes all those solutions which exclude D-1. We will denote the exclusion of D-1 as $\widetilde{\text{D-1}}$. Next we must compute the lower bound for each of these branches.

Lower bounds for the first two branches First we compute the lower bound for the branch which includes D-1. To accomplish this we eliminate row D and column 1 from the matrix. This is shown in Fig. 8-11. Then column reductions are made, followed by row reductions. The columns can be reduced by 3 and 1, and the new matrix reduced by columns is shown in Fig. 8-12. Notice that no row reductions are possible. Therefore Fig. 8-12 represents the new, fully reduced matrix, and it has a lower bound of 63 + 4 = 67. We can therefore conclude that all solutions which include D-1 as an assignment have a minimum cost of at least 67. This lower bound is entered in Fig. 8-9.

Now the lower bound associated with excluding D-1 from all solutions must be determined. To accomplish this we must return to Fig. 8-7. In cell D-1 we enter a dash, which will imply an infinitely high cost if that assignment is made. This is shown in Fig. 8-13. Next, column and row reductions must be taken. From Fig. 8-13 it can be seen that a column reduction of 4 can be taken in the first column. The column-reduced matrix is shown in Fig. 8-14, and we can see that no row reductions are possible. The lower bound for all solutions which exclude D-1 is therefore 63 + 4 = 67. It is entered in Fig. 8-9.

Location

		3	4	5
	3	0	0	1
B	6	2	0	4
	3	0	0	1
	5	1	0	4

Fig. 8-11

Location

		3	4	5
	0	0	0	0
B	3	2	0	3
	0	0	0	0
	2	1	0	3

Fig. 8-12 Fully reduced matrix D-1.

Location

Facility			3	4	5
	4	3	0	0	1
B	5	6	2	0	4
C	4	3	0	0	1
D	–	0	2	1	0
	6	5	1	0	4

Fig. 8-13

Location

	1	2	3	4	5
	0	3	0	0	1
B	1	6	2	0	4
C	0	3	0	0	1
D	–	0	2	1	0
E	2	5	1	0	4

Fig. 8-14 Fully reduced matrix $\widetilde{\text{D-1}}$.

Location

		3	4	5
	0+0=0 0	0+0=0 0	0+0=0 0	0+0=0 0
	3	2	0+2=2 0	3
C	0+0=0 0	0+0=0 0	0+0=0 0	0+0=0 0
	2	1	0+1=1 0	3

Fig. 8-15

At this point we are ready to branch from either D-1 or $\widetilde{\text{D-1}}$. Both have the same lower bound (by coincidence only); but since D-1 has one more assignment than does $\widetilde{\text{D-1}}$, we shall elect to branch from it. We might have to return to $\widetilde{\text{D-1}}$ later to ensure that it does not lead to a better solution, but for now we will neglect it.

Branching from D-1 To branch from D-1, we return to its fully reduced matrix in Fig. 8-12. Again the opportunity costs are computed for each zero cell. It can be seen from Fig. 8-15 that the cell with the highest opportunity cost is cell B-4. Therefore it will serve as the basis for partitioning this reduced problem even further.

We can now enter the second two branches in Fig. 8-9. First we have D-1,B-4, or all of those solutions which include *both* D-1 and B-4 as assignments. Second we have D-1,$\widetilde{\text{B-4}}$, or all of those solutions which include D-1 and not B-4. Next we must compute the lower bounds for these branches.

First we will compute the lower bound for D-1,B-4. Reflecting the fact that B-4 is also assigned, we remove row B and column 4 from Fig. 8-15. The new matrix is shown in Fig. 8-16. Next, column and row reductions are made. These total 1; therefore the lowest bound for this branch is 67 + 1 = 68. It is entered in Fig. 8-9. The new, fully reduced matrix is shown in Fig. 8-17.

Now the lower bound for D-1,$\widetilde{\text{B-4}}$ must be computed. To accomplish this we must return to Fig. 8-12 and enter a dash into B-4. This is shown in Fig. 8-18. It can be seen that no column reduction, but a row reduction of 2, can be made. The fully reduced matrix is shown in Fig. 8-19. Its lowest bound is 67 + 2 = 69. Finally we enter this into Fig. 8-9.

Location

Facility		3	5
A	0	0	0
C	0	0	0
E	2	1	3

Fig. 8-16

Location

Facility		3	5
A	0	0	0
C	0	0	0
E	1	0	2

Fig. 8-17 Fully reduced matrix D-1,B-4.

Location

		3	4	
A	0	0	0	0
B	3	2	–	3
C	0	0	0	0
E	2	1	0	3

Fig. 8-18

Location

Facility		3	4	5
	0	0	0	0
B	1	0	–	1
C	0	0	0	0
E	2	1	0	3

Fig. 8-19 Fully reduced matrix D-1,$\widetilde{\text{B-4}}$.

Where to go next By referring to Fig. 8-9 we can see that at this stage branching can take place from one of three points. These include $\widetilde{\text{D-1}}$; D-1,$\widetilde{\text{B-4}}$; and D-1,B-4. Choosing the alternative with the lowest bound, we therefore elect to branch from $\widetilde{\text{D-1}}$.

Branching from $\widetilde{\text{D-1}}$ To branch from $\widetilde{\text{D-1}}$ we must return to its fully reduced matrix shown in Fig. 8-14. In Fig. 8-20 the opportunity costs are shown in the zero cells. It can be seen that D-2 will be the assignment which will serve as the basis for branching. We therefore branch in the following two directions: $\widetilde{\text{D-1}}$,D-2, and $\widetilde{\text{D-1}}$,$\widetilde{\text{D-2}}$. These are entered in Fig. 8-9.

To find the lowest bound of D-1,$\widetilde{\text{D-2}}$, first we eliminate row D and column 2 from Fig. 8-20. This new matrix is shown in Fig. 8-21. Then we reduce by columns and rows. The total reduction possible is 1; therefore, the lowest bound is 67 + 1 = 68. It is entered in Fig. 8-9. The fully reduced matrix is shown in Fig. 8-22.

Location

Facility

			3	4	5
	0+0= 0 0	3	0+0= 0 0	0+0= 0 0	1
B	1	6	2	0+1= 1 0	4
C	0 0	3	0+0= 0 0	0+0= 0 0	1
D	–	3+0= 3 0	2	1	1+0= 1 0
E	2	5	1	0+1= 1 0	4

Fig. 8-20 Branching from $\widetilde{\text{D-1}}$.

Location

	1	3	4	
A	0	0	0	1
B	1	2	0	4
C	0	0	0	1
E	2	1	0	4

Fig. 8-21

Location

Facility

	1	3	4	5
	0	0	0	0
B	1	2	0	3
C	0	0	0	0
	2	1	0	3

Fig. 8-22 Fully reduced matrix $\widetilde{\text{D-1}}$,D-2.

Now we look at the other branch, $\widetilde{\text{D-1}}$,$\widetilde{\text{D-2}}$. To find its lower bound we must return to Fig. 8-14 and enter a dash in cell D-2. This is done in Fig. 8-23. Next we make column and row reductions. These total 3; therefore the lowest bound of the branch $\widetilde{\text{D-1}}$,$\widetilde{\text{D-2}}$ is $67+3=70$. This is entered in Fig. 8-9. The fully reduced matrix is shown in Fig. 8-24.

Where to go next Referring to Fig. 8-9, it can be seen that there is a tie between the lower bounds of $\widetilde{\text{D-1}}$,D-2 and D-1,B-4. However, since D-1,B-4 has one more assignment than does $\widetilde{\text{D-1}}$,D-2, we will choose to branch from it.

Branching from D-1,B-4 To branch from D-1,B-4 we return to Fig. 8-17. In the zero cells are entered the opportunity costs. This is shown in Fig. 8-25. Cell E-3 will serve as the basis for branching. We can therefore branch in the following two directions: D-1,B-4,E-3 and D-1,B-4,$\widetilde{\text{E-3}}$. These are entered in Fig. 8-9.

To find the lowest bound of D-1,B-4,E-3, first we eliminate row E and column 3 from Fig. 8-25. The new matrix is shown in Fig. 8-26.

Location

	1	2	3	4	5
A	0	3	0	0	1
B	1	6	2	0	4
C	0	3	0	0	1
D	–	–	2	1	0
E	2	5	1	0	4

Fig. 8-23

Location

Facility			3	4	5
A	0	0	0	0	1
B	1	3	2	0	4
C	0	0	0	0	1
D	–	–	2	1	0
E	2	2	1	0	4

Fig. 8-24 Fully reduced matrix $\widetilde{\text{D-1}}$,$\widetilde{\text{D-2}}$.

0+0= 0 0	0+0= 0 0	0+0= 0 0
0+0= 0 0	0+0= 0 0	0+0= 0 0
1	0+1= 1 0	2

Fig. 8-25

It is clear that no column and row reductions are possible. Therefore the lowest bound for D-1,B-4,E-3 is 68 + 0 = 68.

Following our usual pattern, we should now determine the lowest bound of D-1,B-4,$\widetilde{\text{E-3}}$. However, its lowest bound can certainly be no lower than the lower bound of D-1,B-4, the point from which it branches. Therefore at best it would be 68, and this would be the same as D-1,B-4,E-3. We therefore can temporarily avoid evaluating this branch and proceed to branch from D-1,B-4,E-3.

Branching from D-1,B-4,E-3 We return to Fig. 8-26 and compute the opportunity costs. These are shown in Fig. 8-27.

It should be clear that it does not matter which assignment is made. Let us make the assignment in A-2. We can then branch in the following two directions: D-1,B-4,E-3,A-2 and D-1,B-4,E-3,$\widetilde{\text{A-2}}$. These are entered in Fig. 8-9.

Calculating the lower bound for D-1,B-4,E-3,A-2, we eliminate row A and column 2 from Fig. 8-27. Since no column and row reductions can be taken, the lowest bound of this branch is 68 + 0 = 68. This is entered in Fig. 8-9. Again it is not necessary to evaluate D-1,B-4,E-3,$\widetilde{\text{A-2}}$, since its lowest bound will be at least 68.

Solution We now have only one choice left, and that is to assign C to 5. This branch will be depicted as D-1,B-4,E-3,A-2,C-5, and it will have a lower bound of 68 since there is nothing left to reduce. We have a solution!

Notice also that all other branches have a lower bound at least as great as our solution. Therefore it is unnecessary to go back and continue branching, since solutions can be no better than the one we have just computed. If, on the other hand, some of the branches had a lower bound less than 68, we would have to go back to them and branch until the lower bound increased to 68 or we reached a better solution.

$$\begin{array}{c|c|c|}
 & \multicolumn{2}{c}{\text{Location}} \\
 & 2 & 5 \\
\hline
\text{A} & 0 & 0 \\
\hline
 & 0 & 0 \\
\hline
\end{array}$$

Fig. 8-26 Fully reduced matrix D-1,B-4,E-3.

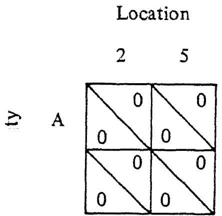

Location
2
5
A
0
0
0
0
0
0
0
0

Fig. 8-27

Computing the cost of our solution The cost of this solution can be computed by referring back to the original problem matrix in Fig. 8-5.

Assignment	*Cost*
D-1	5
B-4	24
E-3	14
A-2	11
C-5	14
Total cost	68 (that is, $68,000)

Notice that it is the same as the lowest bound of our solution branch.

Case study: Glover Landing Glover Landing is a very exclusive apartment complex located on the shores of Marblehead Harbor, approximately 18 miles north of Boston. When it was built in 1966, it won architectural acclaim for its design and for the way in which it harmoniously blended with the natural beauty of the harbor.

It was originally operated as a rental complex, but recently management has decided to sell all 116 apartments as condominiums. As of now, three-quarters of them have been sold.

Parking has always been a problem. The location of its 136 parking places played a secondary role in the design of the complex. Some apartments located on Skinners Head, in fact, are quite far from the nearest parking. As a result it has never been quite clear who should park where. Since it is a town law that each apartment must have its own parking place, management is now faced with making permanent assignments.

If the present parking assignments are made permanent, the current owners definitely have the advantage over those who will eventually purchase the remaining units. New owners will have to walk relatively long distances. Management has therefore decided to review the situation and make new permanent assignments. The question is, Which parking place should be assigned to each apartment?

To help solve this problem, management brought in a consulting firm. The consultant first determined that the objective should be to minimize the distance walked for all tenants. He then decided that this was actually an assignment problem, and that if he could measure all of the distances between apartments and parking places, he could utilize a standard assignment computer code to solve the matrix.

PROCESS LAYOUT: DEPENDENCE

In the last section we studied the case where flow was strictly between a central point and one location. The cost of a particular assignment was determined

From	To t			
	0	5	7	8
	5	0	9	6
3	7	9	0	3
	8	6	3	0

Fig. 8-28 Distance matrix.

To j			
0	5	9	3
4	0	4	2
3	6	0	1
2	7	4	0

Fig. 8-29 Traffic-intensity matrix.

only by the assignment of a facility to that location; it was independent of the assignments made to other locations. It frequently occurs, however, that the cost of assigning a facility to a location is dependent on the locations assigned to other facilities. As an example of this, consider the location of a lathe in a machine shop. If the consequence of the location of this lathe was truly independent of other machine locations, then we would expect that *any* spatial arrangement of the remaining machines in the shop would not affect the cost of the lathe assignment. This, however, is quite unusual. More likely there are recurring patterns of flow between machines. If several of the jobs which are scheduled on the lathe will also be processed by a drill press and a grinder, the location of these machines should be taken into consideration when the lathe is located in the shop. We then say that the consequence of the location of the lathe is dependent on the location of other machines. The magnitude of this facilities location problem is many times that of the independent case.

CRITERION FOR EVALUATING ALTERNATIVE LAYOUTS

Distance and traffic For any given assignment of facilities to locations, its cost is frequently measured by the product of the distance between these locations and the intensity of the traffic flow between these facilities. The distance data should be easy to obtain. The intensity data would either have to be collected over a period of time, uncovered from company records, or estimated.

As an example, consider the problem of assigning four facilities, A, B, C, and D, to four locations, 1, 2, 3, and 4. The distance and traffic-intensity matrices are shown in Figs. 8-28 and 8-29. The distance matrix is symmetrical. That is, the distance from location k to location t is equal to the distance from location t to location k. For example, the distance from 1 to 2 is 5 and the distance from 2 to 1 is also 5. The diagonal is, of course, all zeros. The traffic-intensity matrix, however, is not symmetrical. The intensity of the traffic from machine i to machine j is very infrequently the same as the traffic from j to i. In Fig. 8-29 it can be seen, for example, that the number of trips from

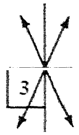


Fig. 8-30 Trial assignment.

machine A to machine C was 9, whereas the number of trips from machine C to machine A was 3.

The evaluation of a layout strategy Let us examine the procedure for determining the cost of a possible layout strategy. Consider the layout shown in Fig. 8-30. Facility A is assigned to location 1, B to 2, C to 4, and D to 3. First the flow between 1 and 2 will be examined. The distance between these locations is 5. The intensity of the traffic between facilities A and B is 5 in the direction A to B and 4 in the direction B to A, for a total intensity of $5 + 4 = 9$. Measuring cost as the product of distance and intensity, the cost of assigning facilities A and B to locations 1 and 2 is $5 \times (5 + 4) = 45$.

Next we turn to the cost of assigning facilities B and C to locations 2 and 4. The distance between these locations is 6 and the intensity is $(4 + 6)$. Therefore the cost is $6 \times (4 + 6) = 60$. We continue to evaluate the cost for all possible flows. In total there are six possible directions of flow. These are all shown in Fig. 8-30. The cost for each of these directions, as well as the total cost of this trial solution, is given in Table 8-1.

All of this effort was necessary to calculate the total cost of but one of the

$$N! = 4! = 24$$

possible assignment strategies. As the number of facilities and locations grows, so does the task of evaluating each possible strategy. For example, if there had

Table 8-1 Cost of trial assignment

Assignment	*Distance × Intensity*	*Cost*
1-2, A-B	$5 \times (5 + 4) =$	45
2-4, B-C	$6 \times (4 + 6) =$	60
4-3, C-D	$3 \times (1 + 4) =$	15
3-1, D-A	$7 \times (2 + 3) =$	35
1-4, A-C	$8 \times (9 + 3) =$	96
2-3, B-D	$9 \times (2 + 7) =$	81
		332

been five facilities to assign to each of five locations, then there would have been $C_2^5 = \frac{5!}{2!3!} = 10$ directions, for which the cost would have to be determined for each of the $5! = 120$ possible strategies. Table 8-1 would then have had 10 rows. Recall that in the independent case, a problem with five facilities and five locations requires only that five numbers be added to establish the total cost of a trial solution. You should now be convinced that the dependent case is much more complex.

Now we turn to ways in which we can solve this problem. We can formulate it as a quadratic programming problem; however, the size of the problem gets quickly out of hand. For those who are interested, the quadratic programming model is presented in the next section as optional material. As you have found to be a common practice, when things get tough, we resort to heuristics. This time is no exception. In the section after the next, you will find a heuristic model for finding a satisfactory solution to the facilities layout problem under dependent conditions.

//// A MATHEMATICAL PROGRAMMING FORMULATION OF THE PROBLEM

We can formulate the dependent case of the facilities layout problem in the following way:

$$\text{Let } X_{ik} = \begin{cases} 1 & \text{if facility } i \text{ is assigned to location } k \\ 0 & \text{otherwise} \end{cases}$$

$$\text{and } X_{jt} = \begin{array}{ll} & \text{if facility } j \text{ is assigned to location } t \\ 0 & \text{otherwise} \end{array}$$

The objective function can then be written as:

$$\text{Minimize } C = \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{t=1}^{n} \sum_{k=1}^{n} D_{kt} I_{ij} X_{ik} X_{jt} \qquad \begin{array}{c} k \neq t \\ i \neq j \end{array}$$

where D_{kt} = distance from location k to t
I_{ij} = traffic intensity between facilities i and j, and *only* in direction i to j

Perhaps some explanation will make this less formidable. First, this objective function essentially considers all possible combinations of assignments. Consider the following simple problem where two facilities must be assigned to two locations.

	To t	
	1	2
From	–	5
	5	–

	To j	
	A	B
From	–	3
	6	–

The objective function can be written in the following way:

$$\text{Minimize } C = D_{21}I_{AB}X_{A2}X_{B1} + D_{12}I_{AB}X_{A1}X_{B2} + D_{21}I_{BA}X_{B2}X_{A1} + D_{12}I_{BA}X_{B1}X_{A2}$$

By substituting the numbers and grouping like terms, we have

$$\text{Minimize } C = 45X_{A2}X_{B1} + 45X_{A1}X_{B2}$$

In this trivial problem we had four terms in the original objective function. As the size of the problem grows, however, the number of terms in the objective function rapidly increases. For example, in a problem with three facilities and three locations, there are 36 terms in the objective function. Can you verify this? As you can see, the optimal solution of large problems is a task of considerable magnitude. This is why heuristic solutions have frequently been used. Now let us turn to the constraints.

The first set of constraints ensures that each location will be assigned but one facility.

$$\sum_{i=1}^{n} X_{ik} = 1 \qquad k = 1, 2, \ldots, n$$

The second set of constraints ensures that each facility will be assigned to only one of n locations.

$$\sum_{k=1}^{n} X_{ik} = 1 \qquad i = 1, 2, \ldots, n$$

In our simple example we would have:

$$X_{A1} + X_{B1} = 1$$
$$X_{A2} + X_{B2} = 1$$

and

$$X_{A1} + X_{A2} = 1$$
$$X_{B1} + X_{B2} = 1$$

Have you discovered why this is a nonlinear programming model? The constraints are indeed linear, but look once again at the objective function. The product $X_{ik}X_{jt}$ is certainly not linear. A program of this type is called *quadratic,* and unfortunately no easy solution technique exists. There are, however, quadratic assignment algorithms that exploit the special structure of this problem.[2] They, too, require considerable computational effort, but using them is much more efficient than solving this as a full-blown quadratic programming problem. ////

---

[2] Eugene L. Lawler, "The Quadratic Assignment Problem," *Management Science,* vol. 9, no. 4, July 1963, pp. 586–599.

A HEURISTIC MODEL

There are several heuristic rules available for solving the dependent-facilities layout problem. The most popular one, CRAFT, has been developed by Buffa.[3] It proceeds by requiring as input an initial trial solution. Either the present layout or a good guess can be supplied. Together with this trial assignment of facilities to locations, a traffic-intensity matrix is required. The heuristics' first step is to compute the cost of this trial solution. It then exchanges pairs of facilities and computes the cost of each new layout. When completed, these new layouts are compared with the original one. If one of them yields a lower cost, it becomes the layout which acts as the standard of comparison for a new round of paired exchanges. The process continues until no further improvement can be found. Now we turn to an example.

First set of paired exchanges Consider the problem shown in Figs. 8-28 and 8-29. Let us use as our input the trial solution illustrated in Fig. 8-30. We record it together with its cost of 332 in Fig. 8-31. Now we can proceed to exchange pairs of facilities. This is accomplished by exchanging the first facility with each of the others. First facility A is exchanged with facility B. This represents trial 2. Then facility A is exchanged with D, and finally it is exchanged with C. These steps are recorded in Fig. 8-31. Each one represents a trial solution whose costs must be computed.

Cost matrix To make the job of determining costs easier, it is helpful to have a cost matrix. It can be compiled by listing all possible pairs of locations down the first column and all possible pairs of facilities across the top row. This is shown in Fig. 8-32. In Table 8-1 the cost of six of these combinations has already been determined. These are entered in the appropriate cells of the cost matrix. The remaining cells can be filled only after the cost for each particular combination has been calculated. Before moving on, you should verify a few of the cell entries.

Evaluation of paired exchanges Now the cost of each remaining trial solution in Fig. 8-31 must be evaluated. For trial 2, the layout is illustrated in Fig.

---

[3] Elwood S. Buffa, Gordon C. Armour, and Thomas E. Vollman, "Allocating Facilities with CRAFT," *Harvard Business Review,* March–April 1964.

	[1	*2*	*3*	*4]*	*Cost*	
Trial 1	A	B	D	C	332	⟵ initial trial solution
Trial 2	B	A	D	C	320	
Trial 3	D	B	A	C	297	⟵ lowest cost
Trial 4	C	B	D	A	331	

Fig. 8-31 First set of paired exchanges.

Table 8-2

Assignment	*Cost*
1-2, B-A	45
2-4, A-C	72
4-3, C-D	15
3-1, D-B	63
1-4, B-C	80
2-3, A-D	45
	320

8-33 and the cost computed in Table 8-2. The cost of this assignment is 320. It is entered in Fig. 8-31. The costs for the other trials are computed and entered in Fig. 8-31. It can be seen that trial 3 is the assignment with the lowest cost. Trial 3 therefore becomes the starting point for a new round of paired exchanges.

Second set of paired exchanges The process is repeated in Fig. 8-34. This time the first trial solution is the best that was uncovered in the last round. Again

	Facility pairs					
Location pairs	A-B	A-C	A-D	B-C	B-D	C-D
1-2	45	60	25	50	45	25
1-3	63	84	35	70	63	35
1-4	72	96	40	80	72	40
2-3	81	108	45	90	81	45
2-4	54	72	30	60	54	30
3-4	27	36	15	30	27	15

Fig. 8-32 Cost matrix.

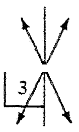


Fig. 8-33 Direction of flows in trial 2.

	[*1*	*2*	*3*	*4*]	*Cost*	
Trial 3	D	B	A	C	297	←
Trial 5	B	D	A	C	299	
Trial 6	A	B	D	C	332	
Trial 7	C	B	A	D	324	

Fig. 8-34 Second set of paired exchanges.

the first facility is exchanged with the other facilities. Then the cost for each of the new trials is computed.

Solution It can be seen that these exchanges did not lead to a cost lower than 297. As a result we have finished our search and choose trial 3 as our solution. If, of course, any other trial solution had displayed a cost lower than 297, the process would have continued for at least another round. Only when the trial which was best in the previous step is also best in the current step is the process complete.

It should be clear that this heuristic does not guarantee the optimal answer. In this example only eight out of the possible $4! = 24$ possibilities were examined. Nevertheless, the heuristic has been found to be quite acceptable in practice.

Comparison of craft with other heuristics The performance of several layout heuristics, including CRAFT, has been compared, using a set of representative layout problems.[4] The objective of the study was to compare these heuristics on the basis of computer-time requirements and computed cost of layout solution (objective function value). The best performance on both counts was achieved by using the CRAFT routine.

Sensitivity Vollman and Buffa[5] have studied the relationship between informal graphical solutions to the layout problem and this heuristic solution. To determine the potential contribution of the heuristic, the standard deviation of the traffic-intensity matrix is computed by regarding each element within the matrix as an observation. Then its coefficient of variation is computed by dividing the standard deviation by the mean. Large coefficients of variation indicate dominance. They found that where dominance is significant, informal methods such as graphical analysis can lead to solutions nearly as good as solutions by the heuristic rule. For example, in one problem where the coefficient of variation was 135, graphical techniques were used to construct a layout. The strategy used was to focus on the dominant flows. The cost of the resulting layout was 12 percent higher than the layout generated by the heuristic rule. In another problem, with a coefficient of variation of 519, the heuristic rule could do only 3 percent better than the informal graphical approach. Their conclusion was that when the coefficient of variation exceeds 200, the problem is dominated by certain flows and informal methods can be used to achieve a layout that will be within 5 percent of the heuristic's solution. Only when dominance is not significant is it worthwhile to use the heuristic rule.

---

[4] Larry P. Ritzman, "The Efficiency of Computer Algorithms for Plant Layout," *Management Science*, vol. 18, no. 5, January 1972, p. 240.

[5] Thomas E. Vollmann and Elwood S. Buffa, "The Facilities Layout Problem in Perspective," *Management Science*, vol. 12, no. 10, June 1966, pp. B450–B468.

SUMMARY

In this chapter we have examined the problems associated with the layout of facilities. The simplest case occurs when the consequence of locating a facility is completely independent of the location of other facilities. To solve this problem the assignment algorithm can be used.

If, however, the consequence of the assignment of a facility to a location depends upon where the other facilities are located, the solution of the problem becomes much more complex. A quadratic programming model was presented which can be used to find the optimal solution; however, it usually becomes too large and unwieldy to be useful. Practitioners have preferred the use of heuristics, the most popular of which is Buffa's CRAFT. However, it finds good, not necessarily the best, answers. These practitioners have found it, nonetheless, quite adequate.

In both cases the major assumptions were that the intensity of traffic was known with certainty and that the cost of an assignment was linearly related to distance. It might be, however, that costs are not linear but behave as a step function. For example, costs could be constant up to a certain volume of work, and beyond that point additional men and equipment would be required. Just after that addition, a considerable amount of idle capacity would exist and small increments to volume would incur little if any extra costs.

The other assumption was that the traffic intensity was known with certainty. In fact the usual approach is to collect traffic-intensity data from company records and use them to estimate the intensity over the planning period. The relevant question that should be asked is, What effect would uncertainty in this traffic-intensity matrix have on the final layout solution? If the final solution is quite sensitive to these data, perhaps a forecast of traffic intensity should be carefully compiled.

Finally, we assumed that the traffic mix and total volume would not change over time. That is, if projects A, B, and C flowed through the shop today, they would flow through the shop tomorrow and the day after in the same ratio and at the same level of intensity. An interesting problem which we have left unanswered is how one would respond to a changing mix and a changing level.

QUESTIONS

Q8-1. Contrast product and process layout. What criteria should be used to choose between them?

Q8-2. What is the difference between the dependent and independent layout problems?

Q8-3. Why is the dependent facilities layout problem more complex than the independent?

Q8-4. Is the facilities layout problem in a hospital any different from that in a manufacturing firm? What about the criterion by which alternative layouts are compared?

Q8-5. What difference in objectives exists between the layout of a manufacturing facility and the layout of a supermarket?

Q8-6. If in the first round of column and row reductions the zeros were situated in such a way that a feasible solution to the problem could be identified, would this be the optimal one? Explain.

Q8-7. Did the strategy taken by the consultant in the Glover Landing case seem reasonable?

Q8-8. Develop a heuristic rule which might be used in place of the assignment algorithm for the Glover Landing case.

Q8-9. Why is Buffa's CRAFT routine a heuristic?

Q8-10. What is the consequence of a matrix which has no clear dominance?

Q8-11. What relationship is there between facilities layout, human behavior, and productivity?

Q8-12. A large Eastern university has announced that its new 15,000-student campus will be subdivided into five campuses with 3,000 students each. Only a very few facilities, such as a central library and a student health center, will be shared by all. What advantages and disadvantages do you see in this layout strategy?

Q8-13. When can the linear assignment algorithm be used to solve integer programming problems?

PROBLEMS

P8-1. Solve the following assignment problem.

		Location			
		1	2	3	
Facility		11	7	4	8
	B	3	14	12	6
	C	9	7	13	10
	D	6	11	5	14

P8-2. The following five facilities must be assigned to five locations. Which strategy will minimize costs?

Location

Facility	1	2	3	4	5
A	6	3	4	2	4
B	7	5	9	2	5
C	4	3	6	8	4
D	3	2	4	3	6
E	6	1	10	7	3

P8-3. The following table represents the cost of assigning facility *i* to location *j*. Assign each of the five facilities in such a way as to minimize costs.

Location

Facility			3	4	
A	12	7	13	6	4
B	9	2	6	11	17
C	5	7	8	7	5
D	12	13	9	9	6
E	6	5	8	8	5

P8-4. Using the heuristic rule and the following initial layout, find a solution to the following problem.

To *t*

Fro	1			
	–	3	6	4
	3	–	9	7
	6	9	–	5
	4	7	5	–

Distance matrix

To *j*

From *i*		B	C	D
A	–	10	7	4
B	6	–	8	6
C	4	8	–	11
D	5	9	4	–

Traffic intensity matrix

Initial layout

[1 4]

B D

P8-5. A shop with four machines has assigned them to the following locations.

Location

Facility	1	2	3	4
	1			
B			1	
C		1		
D				1

Given the following distance and traffic data, determine whether this assignment is the best possible one. If not, find a better solution.

To *t*

From	1	2	3	4
	–	3	7	2
	3	–	6	4
	7	6	–	8
	2	4	8	–

Distance matrix

To

From *i*	A	B	C	D
A	–	4	6	7
B	3	–	5	2
C	6	12	–	7
D	4	7	3	–

Traffic intensity matrix

P8-6. Given the following initial trial solution, find a better layout, using the heuristic rule.

To *t*

Fron			
	–	5	3
2	5	–	2
	3	2	–

Distance matrix

To *j*

From *i*	A	B	C
A	–	3	4.
B	7	–	5.
C	6	4	–

Traffic matrix

[1 3]

B

P8-7. Set up the following quadratic programming model.

To *i*

	1	2
1		
2	6	

Distance matrix

To *j*

	A	B
1	–	7
2	4	–

Traffic matrix

BIBLIOGRAPHY

Buffa, Elwood S.: *Production-Inventory Systems: Planning and Control,* Richard D. Irwin, Inc., Homewood, Ill., 1968, chap. 10.

———, G. C. Armour, and T. E. Vollman: "Allocating Facilities with CRAFT," *Harvard Business Review,* March–April, 1964.

Groff, Gene K., and J. F. Muth: *Operations Management: Analysis for Decisions,* Richard D. Irwin, Inc., Homewood, Ill., 1972, chap. 7.

Lee, R. C., and J. M. Moore: "CORELAP—Computerized Relationship Layout Planning," *Journal of Industrial Engineering,* vol. 18, no. 3, March 1967, pp. 195–200.

Ritzman, Larry P.: "The Efficiency of Computer Algorithms for Plant Layout," *Management Science,* vol. 18, no. 5, January 1972, p. 240.

Vollman, T. E., and E. S. Buffa: "The Facilities Layout Problem in Perspective," *Management Science,* vol. 12, no. 10, June 1966, pp. 450–468.

Wagner, Harvey M.: *Principles of Operations Research,* Prentice-Hall, Inc., Englewood Cliffs, N.J., 1969.

9

Facilities Location

INTRODUCTION

The impetus to embark on a facilities location study can usually be attributed to one or a combination of several reasons. Perhaps an increase in product demand necessitates an addition to the firm's productive capacity. In that case either the present facilities can be expanded, a new plant erected, or the increase subcontracted. Perhaps the location problem must be considered because the market for the firm's product has shifted, or because a new product is being introduced. A change in the cost and availability of input factors is yet another reason why new facilities might be considered. Whatever the reason, once this need is apparent, alternative sites must be identified and both the tangible and intangible costs evaluated. The site which minimizes the sum of these costs is considered the most attractive.

Some of these costs vary with the location chosen, while others do not. In this chapter we are interested only in those costs that do so vary. These might include electric power costs, transportation costs, labor costs, land and construction costs, taxes, and so on. In the next section a closer look will be taken at these and other location-dependent costs. Then we will consider a method for comparing alternative sites when a single new facility is being considered and the dominant costs are both tangible and intangible. Next, a model will be developed which is useful in single-location decision situations when the dominant factors are measurable costs. Finally, we turn to multifacility location problems where the dominant factors are measurable costs.

FACTORS WHICH INFLUENCE LOCATION DECISIONS

Zoning regulations can exert a strong influence over the dimensions and structural characteristics of the facility. This, in turn, might require feedback and subsequent changes during the process planning and plant layout stages. Consequently these regulations must be carefully analyzed before the final choice is made.

If the facility requires large amounts of *power,* then its accessibility and cost are quite important. As a result we find the power-consuming industries located in the Pacific Northwest and in the Southeast.

For those firms that depend on *subcontractors,* the availability of qualified subcontractors represents another consideration in the evaluation of alternative sites.

Shipping and *receiving* facilities are other factors. If air transportation is important, how far away is the airport? Are there frequent flights to major market areas? If trucks will be used to ship the product, will the local facilities be adequate? Perhaps shipping by ocean, canal, or St. Lawrence Seaway might be efficient. If so, a location near a port narrows the range of alternatives.

In some cases the *cost of transporting* the input factors to the facility or the product to the market varies considerably between location alternatives. When the weight or volume of the product is substantially reduced during the production process, as in the transformation of iron ore into steel, the emphasis will be on locating close to the source of the inputs. In general, as the weight or volume of the product increases in relation to its value, the tendency will be to maintain multiple factories and warehouses. A good example of this is found in the brewing industry.

The availability of an adequate *labor supply* can play a major role in plant location. Textile plants of New England, for example, have abandoned this region for the lower labor costs of the South. In the past the nonunionized labor pool in the South, as a source of cheap labor, has attracted many labor-intensive industries. The market mechanism, albeit slow, responds with an influx of firms to this area. However, as the demand for labor goes up in this region, if there is little or no change in the labor supply, wage rates will increase. The move could therefore prove to be a temporary respite.

In some cases the firm's search for low labor costs has led them to *foreign operations*. It is important, however, not to confuse wage rates and total labor costs. Although some countries entice American manufacturers with low wage rates, the total labor cost finally embodied in the product can be higher than expected because of the differences in efficiency between American and foreign labor.

Frequently foreign operations are a response to competitive pressures in once unchallenged foreign markets. For example, the potential of the Common Market has prompted many firms to establish foreign manufacturing and sales centers. Therefore, in this burgeoning era of the multinational corporation, the process of facility location has broadened to such an extent that it now encompasses the entire free world.

Another factor which can affect the location decision is *land cost.* Its magnitude depends upon whether the facility must be located in an urban, suburban, or rural setting. Telephone switching exchanges, hospitals, post offices, and downtown airline ticket offices must be centrally located, and therefore there is little leeway for them in the choice of a location. This choice, however, is not nearly so limited for most other industries, and the opportunity to compare land prices over a much broader geographic region does exist. In addition to these land costs, *construction* costs warrant close attention. They also vary between location alternatives.

Frequently a locality offers to an attractive firm free land and a substantial *tax* concession for perhaps five years. In many cases the package may not be as attractive as it sounds. In the long run it is impossible for the firm to receive public benefits of sewage disposal, police and fire protection, schools for employees' children, and so on, at no cost. What appears as a temporary advantage may turn out to be a long-run disadvantage as the firm takes over its share of the tax burden once the period of immunity is over.

What must be considered are the long-run as well as the short-run tax implications. The efficiency of the town management, as well as its indebtedness, is a clue which should be examined carefully.

When the firm must locate close to its *market,* the choice of location is severely restricted. Examples include hospitals, schools, churches, telephone switching exchanges, police and fire stations, and cable TV companies.

Recently a new dimension has been added to the facilities location problem: compliance with *pollution* standards. In fact, these new constraints have risen to positions of primary importance in many location decisions.

Since pollution standards are not uniform for all states or within states, there remain significant locational benefits. For heavy users of air and water resources, these could easily prove to be the decisive factor in choosing a new location.

For example, several states have adopted some variation of the following scheme for classifying bodies of water.

Class A: Suitable for any water use, including drinking, recreation, and fish habitat.

Class B: Suitable for bathing and recreation, irrigation, and agricultural uses; good fish habitat; good esthetic value. Acceptable for public water supply with filtration and disinfection.

Class C: Suitable for recreational boating, irrigation of crops not used for consumption without cooking, habitat for wildlife and common food and game fishes indigenous to the region, industrial cooling, and most industrial-process uses.

Class D: Suitable for transportation of sewage and industrial wastes without nuisance, and for power, navigation, and certain industrial uses.

The consequence of this is that the effluent generated from a process would have to be treated according to the river into which it is discharged. For example, the costs of waste treatment for a plant located on a river classified as "A" would be much higher than if the plant were located on a class "D" river.

An air shed into which stack effluents are discharged may also impose standards which are quite different from those of other air sheds. For example, sulfur oxide emission standards in New York City are substantially more severe than those in cities less than 75 miles away.

There are, then, locational advantages. When large quantities of effluent must be disposed into the environment, these advantages become more and more important.

A striking example of this new problem is the location of a power plant. In the past, if you needed additional power, the location of the plant site was not considered the dominant problem. Today, however, it is impossible to locate a plant just anywhere. Public pressure and pollution standards prohibit it. Environmentalists have warned us that the hot water discharged from the plant can seriously upset the balance of nature. Therefore, if this water is discharged into a river, lake, or ocean, the flora and fauna may never be the same again. The power plants have responded with new plant designs that are directed at reducing their impact on the environment. These include cooling towers, holding ponds, and discharge pipes which extend well into the ocean. Stack emissions are another example of effluents that will no longer be tolerated in unlimited quantities. It should be clear to you by now that the power companies face a new problem which frequently dominates the location decision.

Noise pollution is another environmental factor to which a firm must respond. Consider the problem in locating a supplemental airport for a large city, or even in expanding a large metropolitan airport. Public pressure groups are having a good deal of success in making these previously unconsidered noise costs a very real component of the project cost. In some cases, when all of the costs, including these noise costs, are tallied, the benefit of the project is less than its cost.

Since these pollution problems are quite new, and since much time and money may be spent on this problem by the firms of the future, the next chapter has been devoted exclusively to them.

Finally we have *convenience factors.* Although it is unlikely that they will sway a location decision, they are often considered quite important. Included in this category are adequate parking facilities, public transportation, and even athletic facilities for employees.

Case study 1: location of a telephone switching exchange Consider the problem of determining the location of a telephone switching exchange. The function of an exchange is to switch incoming calls to the appropriate subscriber and outgoing calls to the appropriate terminating telephone exchange. Each exchange area serves approximately 10,000 subscribers. To accomplish its function, the exchange must be wired to each telephone instrument. This results in a substantial investment in copper wire. The more wire that has to be run to subscribers, the larger the investment and the maintenance costs. The objective, then, is to find that location within a defined set of telephone subscribers which will minimize the amount of wire required.

Both General Telephone and Electronics Corporation and American Telephone and Telegraph Company have developed computer programs for solving this problem. As input data, the program requires the geographic mix of telephone subscribers. A trial location is made and by means of a heuristic the location is varied until an improvement can no longer be obtained. The solution, then, specifies the best of these possibilities. However, whether or not this location can be purchased is another question. If it can be purchased at a reasonable price, the problem is solved. If not, alternative sites must be evaluated to determine which one of them will come closest.

Case study 2: efficiency of ambulance service The City of New York provides ambulance service to its approximately 10 million inhabitants. In an effort to deliver this service more efficiently, the office of the mayor commissioned their management science team to undertake a rather extensive study of the system.[1]

At the time the study was undertaken, approximately 109 ambulances were stationed at a total of 49 hospitals. The ambulances assigned to a particular hospital responded only to calls which emanated from its well-defined hospital district.

One objective of the study was to determine whether or not the ambulances were located in the best possible locations. If not, what was the best placement? So that alternative placements could be compared, the criterion chosen was response time. This is the period between the receipt of the call at the ambulance station and the arrival of an ambulance at the scene.

The plan for the study was to compare the present system with a set of different strategies in which ambulances would be located at satellite garages. These satellite garages would be positioned at demand centers within the hospital district. The hypothesis was that with the ambulances located directly in areas of high demand, response time would decline.

---

[1] This study was reported in E. S. Savas, "Simulation and Cost-Effectiveness Analysis of New York's Emergency Ambulance Service," *Management Science*, vol. 15, no. 12, August 1969, p. B-608.

The problem then was reduced to one which required the development of a model for evaluating the effect that different locational strategies had on response time. A model of this nature is quite complex, since the level of service depends on at least the following five factors:

1. Geographic distribution of calls throughout the district
2. Frequency of calls
3. Number of ambulances in the district
4. Location of the hospital in the district
5. Location of ambulance garages in the district

The only way to cope with a model which displays this level of complexity is by simulation. This was precisely how they proceeded.

About 175,000 calls were simulated for the Kings County Hospital district in Brooklyn. For the present system which includes seven ambulances stationed at the hospital, the response time which was generated by the simulation analysis was 13.5 minutes. When a satellite was added, the response time slowly declined as more ambulances were shifted to it. Only an 11 percent reduction was accomplished when all of the seven ambulances were assigned to the satellite garage as their new home base of operation. This reduced response time from 13.5 to 12.0 minutes.

Also subjected to the simulation model was the assignment of additional ambulances at the hospital. When their strength was increased from seven to ten, response time dropped only from 13.5 to 13.2 minutes.

Another simulation was run with 10 ambulances, but this time they were distributed between the hospital and the satellite garage. When six were located at the hospital and four at the satellite, response time dropped to 10.9 minutes. This strategy promised the most immediate improvement.

As is so often the case in studies of this nature, the analysis of relatively minor modifications to the system suggests even more beneficial large-scale changes. In this case it became apparent that rather than optimally designing each preconfigured hospital region, it would be more effective to start from scratch and separate the hospital from the ambulance services. Under this radically new approach, the ambulances would be stationed at patient demand centers without regard to where the hospital was located. After the patient was picked up, he would be taken to the closest hospital which could administer treatment.

A central dispatching facility would be required in which up-to-date information on the status of each ambulance and even of each hospital would be available. An integrated computer-based information system could eventually be envisioned where the following sequence of events would be initiated by a call for service.

1. The dispatcher receives the call and assigns the nearest ambulance. By employing statistical techniques the computer could even advise him whether to assign the call immediately to a relatively distant but available ambulance or to wait a short time for another ambulance which is currently on an assignment but close to being available and much nearer to the demand point.
2. After picking up the patient, the ambulance driver determines if hospitalization is necessary and, if so, the nature of the emergency (burn, cardiac, poison, and so on).
3. If a hospital bed is required, the dispatcher can determine the nearest hospital with the appropriate medical facilities and bed.
4. The dispatcher instructs the driver to proceed to that hospital.
5. Dispatcher advises hospital if special aid or equipment should be ready.

Even without this fancy information system, it turns out that separation of ambulance and hospital services can be very effective. A simulation of this strategy was undertaken for the Kings County region. Ten ambulances were scattered at the most active demand centers. Response time dropped 30 percent. This strategy was, then, clearly superior to all of the others.

MEASURING THE RELATIVE MERITS OF SINGLE-FACILITY ALTERNATIVES WHEN THE DOMINANT FACTORS ARE BOTH TANGIBLE AND INTANGIBLE

A complete list of location factors that must be taken into consideration mixes both tangible and intangible factors. Tangible costs may include land costs, building costs, taxes, transportation costs, labor costs, power and water costs, and some pollution prevention costs. Intangible factors may include the political climate, attitudes in the community, future tax burdens, and the reliability of the labor pool.

If everything could be easily measured in dollars, there would be no problem. Unfortunately this is not the case. Not only are some of these factors difficult to measure, but many are measured in different dimensions. How, then, can a figure of merit for each alternative be determined? One way is to utilize a technique called *dimensional analysis*.

Consider a problem in which a single facility must be located. The choice has been narrowed down to two alternatives: location 1 in Chicago and location 2 in Dallas. The following six factors are seen as the ones which should guide the decision: cost of the land, cost of the building, labor supply, taxes, community attitudes, and research climate. The land and building costs are estimated in Table 9-1. Taxes are estimated over the life of the plant, and the present value for each alternative is also presented in Table 9-1. The research climate depends upon educational facilities, other research organizations, and local professional societies. The presence of these helps to attract and keep

Table 9-1 Dimensional-analysis example

	Factor	$F_{i,1}$ *Alternative 1*	$F_{i,2}$ *Alternative*	W_i *Weight*
1	Land cost	600,000	300,000	5
2	Building cost	1,000,000	800,000	5
3	Taxes	800,000	500,000	5
4	Labor supply	2	1	3
5	Research climate	1	2	2
6	Community attitudes	3	1	3

Table 9-2 Dimensional-analysis scale

	Points
Research climate	
No schools, research organizations, or societies exist	5
A few low-quality schools but nothing else	4
Good research organizations but nothing else	3
Good schools, research groups, and societies	2
Excellent research climate	1
Labor supply	
None available	5
Limited supply at premium rates	4
Acceptable supply for now but doubtful for future	3
Adequate supply of acceptable ability	2
Plentiful supply of excellent ability	1
Community attitudes	
Active pressure groups against company activity	5
Will accept company activity with reservations	4
Ambivalent	3
Cooperative	2
Very cooperative, encouraging, and helpful	1

good research personnel. To evaluate this and other intangible factors, a five-point scale is developed. The scales for research climate, labor supply, and community attitudes are shown in Table 9-2. Each of these alternatives is scored, and the results are shown in Table 9-1.

Next, the importance of each factor must be weighed. The dollar costs of land, buildings, and taxes are considered the most important factors and are given equal weights of 5. Labor supply and community attitudes are considered equally important and are given a weight of 3. Finally, research climate is given a weight of 2. These weights could be determined either by the individual in charge of the project or by group consensus.

To combine all of these factors, a ratio of the alternative 1 figure to the alternative 2 figure is computed. This in turn is raised to its weighting factor. Finally the product of all ratios is computed.

$$P_{1,2} = \left(\frac{F_{1,1}}{F_{1,2}}\right)^{W_1} \left(\frac{F_{2,1}}{F_{2,2}}\right)^{W_2} \left(\frac{F_{3,1}}{F_{3,2}}\right)^{W_3} \left(\frac{F_{4,1}}{F_{4,2}}\right)^{W_4} \left(\frac{F_{5,1}}{F_{5,2}}\right)^{W_5} \left(\frac{F_{6,1}}{F_{6,2}}\right)^{W_6}$$

where W_i = weight for factor i
$F_{i,1}$ = cost or score for factor i under alternative 1
$F_{i,2}$ = cost or score for factor i under alternative 2

$$P_{1,2} = \left(\frac{600{,}000}{300{,}000}\right)^5 \left(\frac{1{,}000{,}000}{800{,}000}\right)^5 \left(\frac{800{,}000}{500{,}000}\right)^5 \left(\frac{2}{1}\right)^3 \left(\frac{1}{2}\right)^2 \left(\frac{3}{1}\right)^3$$

$$P_{1,2} = 413$$

If the result is greater than 1, alternative 2 is chosen, since the costs in the denominator are less than those in the numerator. If it is less than 1, alternative 1 is preferred. Also note that the product has no dimension. In our example the product is 413, and since this is greater than 1, alternative 2 is chosen.

This analysis is useful when many factors measured in different dimensions must be considered. Next we turn to the case where the predominant factors to be considered can be measured in dollars.

SINGLE-FACILITY LOCATION DECISIONS WHEN THE DOMINANT FACTORS ARE MEASURABLE COSTS

A FIXED- AND VARIABLE-COST MODEL

If one plant must be chosen from among several alternatives, and if most of the important considerations can be measured in dollars, a breakeven analysis can be performed.

First, the fixed and variable costs are identified for each alternative. Figure 9-1 illustrates one such breakdown. Finally, all the total-cost curves are plotted

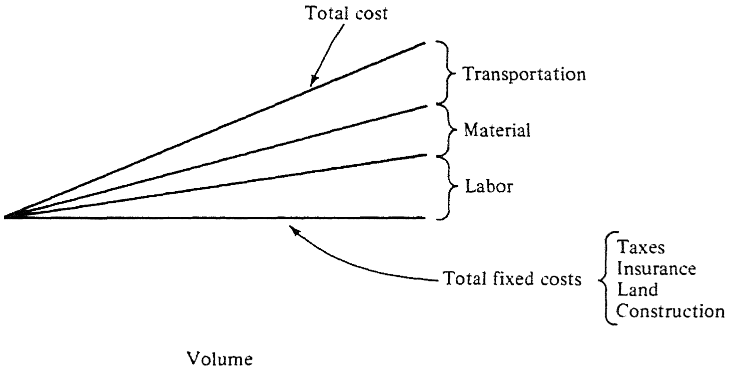


Fig. 9-1 Variable and fixed costs for alternative C.

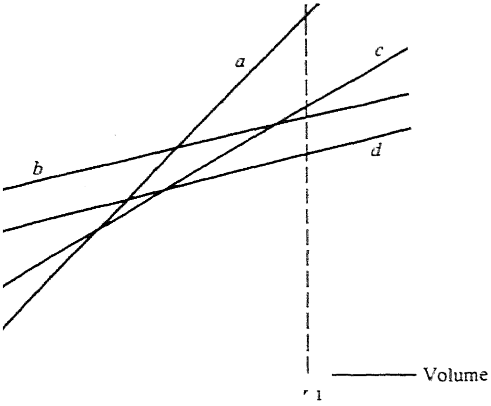


Fig. 9-2 Total-cost curves for each alternative.

on one graph, as shown in Fig. 9-2. Once the level of activity is forecasted, the alternative with the lowest total cost is selected. In this particular example and for an expected volume V_1, alternative *d* is the most economical.

CAPACITY DECISIONS

Another problem that must be considered at this point is the capacity of the facility. Should one be chosen which meets the requirements just for this year, the next 5 years, the next 10? Larger facilities mean a greater investment, but there are several advantages that must also be considered. In many industries the per-unit costs of production, up to a certain level of plant size, decrease with increasing size of the facility. Beyond that point, they increase. This economy of scale can be attributed to the operating efficiency of large plants whose volume is high enough to warrant efficient automatic equipment with product layout formats. Therefore, the construction of a larger plant, which will probably be underutilized in its first few years, may eliminate the necessity for a series of less efficient smaller plants in the future.

Another advantage is that in a period of rising land and construction costs, a larger plant which is constructed early may prove to be a profitable investment.

Given that there are several different plant sites under consideration a present value analysis could be undertaken for each one. Each of the alternatives evaluated would include a certain size plant to be constructed now and any new plants or additions that would be necessary over the planning horizon of the analysis. Finally, the benefit/cost ratio for each of these alternatives would be computed and compared.

MULTIFACILITY LOCATION DECISIONS WHEN THE DOMINANT FACTORS ARE MEASURABLE COSTS

A common problem found in large firms is that there already exists a network of facilities. If a new facility is to be added, the problem is one of determining where it should be located. The consequence of each location alternative will depend on where in relation to existing facilities this new one is placed.

In our first multifacility model only one new addition to an existing complex of facilities can realistically be evaluated at one time. This is so because our model is essentially a variable-cost model. However, the second model, which takes into consideration fixed and variable costs, can cope with any number of additions to an already existing complex of facilities.

A VARIABLE-COST MODEL

The problem Consider the problem faced by the Sudsdown Company, a producer of beer. They currently maintain two production facilities, one in Pittsburgh and one in Denver. From these plants they ship to regional warehouses located in New York City, Chicago, and Los Angeles. Recently the output from their two plants has fallen short of the warehouses' demand. Management sees this as a permanent change in the demand for their beer.

In response to this shortage they have decided to expand their capacity and add a third brewery. An exhaustive search has led to the consideration of two sites: one in Milwaukee and the other in San Francisco.

The warehouse requirements, production capacities, per-unit variable production costs of the existing and proposed facilities, and the building and equipment costs are shown in Table 9-3. Transportation costs from all possible plant sites to warehouses are given in Fig. 9-3. The problem is to select a plant which will minimize the sum of production and distribution costs for the entire network.

At first glance it looks as if the Milwaukee plant, with a per-unit production cost of $.60, is the best choice. It may, however, not be the choice once the interactions with the other plants and warehouses are considered. For example, a plant located in northern Idaho may achieve very low production costs, but when transportation costs to warehouses are considered, it may become quite unattractive. On the other hand, consider a plant with higher-than-average production cost. It may, however, be located in such a way that shipping routes for the entire network of plant and warehouses can be rearranged so that when total transportation costs for the entire system plus the production costs of the new plant are compared with other alternatives, it may be the most economical solution. The conclusion is that any change of the system of plants and warehouses requires analysis of how this change affects the cost of operating the entire system.

		To		
		New York	Chicago	Los Angeles
From	Pittsburgh	4	6	12
	Denver	7	5	6
	Milwaukee	8	2	5
	San Francisco	13	7	3

Fig. 9-3 Transportation cost (dollars per 100 cases).

Table 9-3 Location costs

	Warehouse demand (cases)	
New York	100(00)	
Chicago	130	
Los Angeles	210	
Total	440	
	Production capacity (cases)	**Per-unit (case) production costs**
Pittsburgh	90(00)	$.50
Denver	170	.55
Total	260	
Shortage:	440 − 260 = 180	
	Estimated variable per-unit (case) production costs	
Milwaukee	$.60	
San Francisco	.65	
	Building and equipment costs	
Milwaukee	$13,100 per year over the life of the plant	
San Francisco	12,800	

To investigate the consequence of adding a new plant to the system presented in Table 9-3 and Fig. 9-3, each of the possible alternatives must be analyzed separately. The first step will be to determine the total transportation cost of the system if Milwaukee is chosen. To accomplish this, a transportation strategy must be developed. In Fig. 9-4 the transportation costs between the plants and warehouses are entered in the appropriate cells. The demand gen-

	Warehouse			
	New York	Chicago	Los Angeles	Capacity
Pittsburgh	4	6	12	90
Denver	7	5	6	170
Milwaukee	8	2	5	180
Demand	100	130	210	440

Fig. 9-4 Transportation costs with Milwaukee alternative.

erated by each warehouse is given along the bottom row and the capacity of each plant is given down the last column. The capacity of the Milwaukee plant is chosen to match exactly the needed capacity. As a result the total supply equals the total demand.

The objective is to assign quantities to each plant-warehouse combination in such a way that the transportation costs are minimized and the demand and capacity constraints are not violated. This represents the format of the classical transportation method of linear programming. As you will see later, it is also a special case of linear programming.

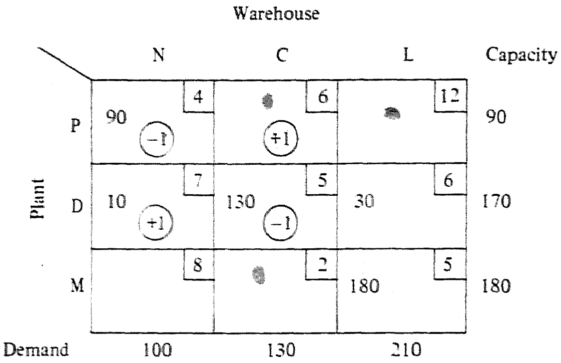
When it can be used, the appeal of the transportation method is the efficiency with which a solution can be reached. The name, *transportation method*, can be traced to its early use for solving transportation-type problems very much like the one on which we are now working. Other problems, not necessarily having to do with transportation but which can be set up in matrix form with a similar constraint structure, can also be solved by this method. Those of you who have been exposed to this algorithm may want to read through the next sections quickly; the others should study these sections carefully.

Transportation method There are several methods for solving the transportation problem, and some are frequently more efficient than others. The method that is chosen here is not very efficient, but it is very intuitive. Since the practitioner will seldom solve a matrix by hand, an understanding of the technique is more important for our purpose than efficiency.

Initial allocation The first step is to make a trial allocation by starting in the upper left-hand corner. In this *northwest corner*, from which comes the name for this method, allocate as much as possible—in our case, 90 units. It is entered in Fig. 9-5. Now, working out from this corner, we assign as much as possible to each cell. The Pittsburgh–Chicago and Pittsburgh–Los Angeles cells cannot be assigned anything, since the capacity constraint for the Pittsburgh plant has already been met. The Denver–New York cell can be assigned but 10 units, since New York demands only 100 units. The Denver-Chicago cell can be assigned 130 units, and this meets the Chicago demand. Then the Den-

	Warehouse			
	New York	Chicago	Los Angeles	Capacity
Pittsburgh	4 90	6	12	90
Denver	7 10	5 130	6 30	170
Milwaukee	8	2	5 180	180
Demand	100	130	210	

Fig. 9-5 Northwest corner allocation.



Plant \ Warehouse	N	C	L	Capacity
P	4 90 (−1)	6 (+1)	12	90
D	7 10 (+1)	5 130 (−1)	6 30	170
M	8	2	5 180	180
Demand	100	130	210	

Fig. 9-6 Evaluation of empty cell P-C.

ver–Los Angeles cell can be assigned 30 units to reach the capacity of the Denver plant. Finally, 180 units can be assigned to Milwaukee–Los Angeles.

It is important at each step that exactly $n + m - 1$ routes be assigned, where n is the number of rows and m is the number of columns in the problem. In our case, $n = 3$ and $m = 3$; therefore, $n + m - 1 = 3 + 3 - 1 = 5$. We see that five routes have indeed been assigned. We can therefore proceed with the solution. When this is not the case and less than $n + m - 1$ routes are assigned, we have a situation of degeneracy. One of the exercises at the end of this chapter gives you the opportunity to solve a problem of this type.

Evaluating empty cells The next step in the solution process is to determine whether a cost savings can be realized by a shift in allocation from a cell already assigned to one which is unassigned. To accomplish this we focus on one unassigned cell at a time. After we have evaluated the consequence of shifting into each of the empty cells, we choose to reassign to that cell which has the largest cost savings. Let us proceed by determining the consequence of moving into the Pittsburgh-Chicago cell.

When evaluating an empty cell, changes must be made in such a way that the constraints along the rim of the matrix are not violated. If one unit is brought into the P-C cell, one unit must be taken out of the D-C cell; otherwise the Chicago demand constrant will be violated. If this one unit is taken out of the D-C cell, one unit must be added to the D-N cell; otherwise the Denver capacity constraint will be violated. Finally, a unit must be taken out of the P-N cell, since a unit was added to the D-N cell. This pattern of exchange is shown in the circles of Fig. 9-6. The net effect to the rim constraints is zero.

Two aspects of this series of moves should be mentioned. First, moves into cells other than the one being evaluated are restricted to cells which already have assignments. For example, if cell P-C is being evaluated, a move of one unit out of cell P-L is not allowed. If it were, one would have trouble interpreting the negative assignment in cell P-L.

The second aspect of the moves made around the matrix really follows from the first. Namely, there is only one possible pattern of moves for the evaluation of each empty cell. Try establishing any other pattern of moves for empty cell P-C. You should find that the one just presented is the only one possible.

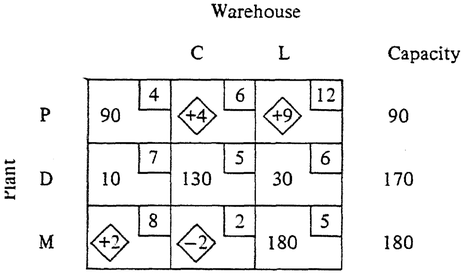
Now that the *pattern* of moves for the evaluation of a cell has been determined, the next step is to compute the *economic consequence* for this series of moves. If a unit is added to P-C, costs increase by 6 for every unit added. For every unit removed from D-C, costs decrease by $5. For every unit added at D-N, costs increase by $7; and for every unit removed from P-N, costs decrease by $4. The net consequence of this shift is $6 - 5 + 7 - 4 = +4$. That is, for every unit moved into P-C by this sequence, net costs increase by $4. Certainly not a profitable reallocation. This per-unit consequence is entered in the P-C cell of Fig. 9-7, so that we can keep track of the profitability associated with each alternative.

Next we will evaluate P-L to determine the consequence of a reallocation into this empty cell. The pattern of moves is into P-L out of D-L into D-N out of P-N. The net per-unit cost consequence is $+12 - 6 + 7 - 4 = +9$. It is entered in the P-L cell of Fig. 9-7.

Next empty cell M-C is evaluated. The pattern of per-unit moves is into M-C out of D-C into D-L out of M-L. The net per-unit cost consequence is $+2 - 5 + 6 - 5 = -2$. Therefore, for every unit moved into M-C there is a net cost *savings* of $2. It is still premature to make any decision. Only after *all* of the empty cells are evaluated can a reallocation be undertaken.

The last empty cell to be evaluated is M-N. The pattern of moves is into M-N out of D-N into D-L out of M-L. The net per-unit cost consequence is $+8 - 7 + 6 - 5 = +2$.

Making the shift A reallocation is made into that empty cell which displays the largest net per-unit cost savings. In Fig. 9-7, cell M-C is that cell. We concluded that a shift into M-C resulted in a net savings of $2 for *each* unit shifted. The next question, then, is, *How many* should be shifted? To answer this, the pattern of moves that established the $2 cost savings must be reconsidered. For every unit added to M-C, a unit was removed from D-C, added to D-L, and removed from M-L. Units were therefore removed from two cells, D-C and M-L. The *maximum number* of units that could be removed is 130, since any number over this would cause a negative allocation in cell D-C. In



Plant		Warehouse			
			C	L	Capacity
P		4 90	6 +4	12 +9	90
D		7 10	5 130	6 30	170
M		8 +2	2 −2	5 180	180
	Demand	100	130	210	

Fig. 9-7

			Warehouse		
			C	L	Capacity
Plant	P	90			90
	D	10		160	170
	M		- 130	5 50	180
Demand		100	130	210	

Fig. 9-8 Solution to Milwaukee alternative.

fact, if more than 180 are removed, negative allocations would occur in cells D-C and M-L.

We therefore add 130 units to cell M-C, remove 130 units from cell D-C, add 130 units to cell D-L, and remove 130 units from cell M-L. In this way we ensure that the rim constraints are not violated. The new matrix is drawn in Fig. 9-8. Recall that the net per-unit consequence of this move was \$2. Therefore, by moving 130 units, a savings of 2 × 130 = \$260 has been effected.

Another iteration At this point the process of evaluating each empty cell is repeated to determine whether further reallocation can reduce costs even more. All of the cells but P-C are straightforward. The pattern of per-unit moves for this cell is into P-C out of M-C into M-L out of D-L into D-N out of P-N. The net per-unit cost consequence of this is $+6-2+5-6+7-4=+6$. This and the other net per-unit cost consequences are entered in Fig. 9-8. Since none of them are negative, it is not profitable to reallocate any further. We have reached the optimal solution.

Solution of the Milwaukee alternative The optimal strategy for the Milwaukee alternative is, therefore, to ship according to the following plan.

Pittsburgh to New York	90
Denver to New York	10
Denver to Los Angeles	160
Milwaukee to Chicago	130
Milwaukee to Los Angeles	50

The transportation cost of this strategy is

$$90(4) + 10(7) + 160(6) + 130(2) + 50(5) = \$1{,}900$$

Solution of the San Francisco alternative Next the optimal transportation strategy for the San Francisco alternative must be determined. Once again, the assignments to the first matrix are made according to the "northwest corner" rule. Then each empty cell is evaluated. It can be seen from Fig. 9-9 that each of the evaluations is positive; therefore it is not profitable to reallo-

		Plant			
		(N) New York	(C) Chicago	(L) Los Angeles	Capacity
Warehouse	Pittsburgh (P)	4 90	6 +4	12 +5	90
	Denver (D)	7 10	5 130	6 30	170
	San Francisco (S)	13 +4	7 +5	3 180	180
	Demand	100	130	210	440

Fig. 9-9 Solution to San Francisco alternative.

cate. The optimal solution for this strategy is to ship according to the following plan.

Pittsburgh to New York	90
Denver to New York	10
Denver to Chicago	130
Denver to Los Angeles	30
San Francisco to Los Angeles	180

The cost of the plan is

$$90(4) + 10(7) + 130(5) + 30(6) + 180(3) = \$1{,}800$$

Comparison of alternatives Now we can compute the building, production, and shipping costs for each strategy. Since the Pittsburgh and Denver plants will operate at full capacity regardless of the alternative chosen, production costs can be ignored. For the strategy which includes the Milwaukee plant we therefore have

$$\begin{aligned} \text{Building and equipment costs} &= \$13{,}100 \\ \text{Production costs (M)}: 18{,}000 \times \$.60 &= 10{,}800 \\ \text{Transportation costs} &= \underline{1{,}900} \\ \text{Total costs} &= \$25{,}800 \end{aligned}$$

The strategy which includes the San Francisco plant has the following cost:

$$\begin{aligned} \text{Building and equipment costs} &= \$12{,}800 \\ \text{Production costs (S)}: 18{,}000 \times \$.65 &= 11{,}700 \\ \text{Transportation costs} &= \underline{1{,}800} \\ \text{Total costs} &= \$26{,}300 \end{aligned}$$

It therefore appears that, from a cost point of view, the Milwaukee plant is the favored alternative.

Linear programming, transportation, and assignment methods We digress for a moment to investigate the relationship between linear programming, the transportation, and the assignment methods.

The transportation problem as depicted in Fig. 9-4 can be written as a linear programming problem. Let X_{ij} be the amount shipped from plant i to warehouse j. Then the objective function can be written as

$$\text{Minimize } C = 4X_{11} + 6X_{12} + 12X_{13} + 7X_{21} + 5X_{22} + 6X_{23} + 8X_{31} + 2X_{32} + 5X_{33}$$

The capacity constraints are

$$X_{11} + X_{12} + X_{13} \leq 90$$
$$X_{21} + X_{22} + X_{23} \leq 170$$
$$X_{31} + X_{32} + X_{33} \leq 180$$

Next we have the demand constraints:

$$X_{11} + X_{21} + X_{31} \geq 100$$
$$X_{12} + X_{22} + X_{32} \geq 130$$
$$X_{13} + X_{23} + X_{33} \geq 210$$

Finally the nonnegativity constraints:

$$\text{All } X_{ij}\text{'s} \geq 0$$

In general notation, we can write this as

$$\text{Minimize } C = \sum_{i=1}^{m} \sum_{j=1}^{n} C_{ij}X_{ij}$$
$$\sum_{j=1}^{n} X_{ij} \leq L_i \qquad i = 1, 2, \ldots, m$$
$$\sum_{i=1}^{m} X_{ij} \geq D_j \qquad j = 1, 2, \ldots, n$$
$$\text{All } X_{ij}\text{'s} \geq 0$$

where C_{ij} = cost of shipping 1 unit from source i to destination j
L_i = capacity of source i
D_j = demand from destination j

Since any problem which fits the transportation format can be written as a linear programming problem, why bother with a transportation algorithm? If you compared the time it would take to solve our example by linear programming and by the transportation method, you would have the answer. The transportation method is much more efficient. Unfortunately we cannot use it for all linear programming problems, since the transportation algorithm can be used only for those problems which can be set up in a matrix as in Fig. 9-4.

Our conclusion, then, is that the transportation method is a special case of linear programming. Therefore, when a problem displays this format, we exploit its structure by utilizing an efficient solution technique.

What about the assignment problem whose format (shown in Fig. 8.1) was much like that of the transportation problem? It, in turn, is a special case of the transportation problem. This stems from the fact that its rim constraints are all equal to 1. Accordingly, the assignment problem is easier yet to solve.

There is another distinct advantage associated with the transportation format and its special cases: the solution to the matrix will always be integer-valued. Therefore, the solution will never suggest that $13\frac{1}{2}$ cases be shipped from plant 1 to warehouse 3. This is an especially valuable aspect, since without it problems which require integer solutions would have to be solved by integer programming methods.

//// A FIXED- AND VARIABLE-COST MODEL

The problem Consider a group of plants serving a number of market areas. In the previous section the concern was with the location of *one* additional plant. Perhaps a more realistic statement of the problem would be, How many plants should be added, and where should they be located? Once this extra dimension is added to the problem, it becomes necessary to consider the fixed costs of setting up a plant as an explicit part of the location model. Otherwise, a plant might be located next to each customer. The trade-off, then, is between increased overhead costs from more plants and reduced transportation cost.

In addition to the fixed cost of opening a plant there is often a fixed cost associated with setting up a route between that plant and the warehouse which it services. For example, when a truck route between two centers is established, there will be a large fixed cost for operating the vehicle. There will also be a smaller variable cost which is related to the volume shipped.

A mixed-integer programming formulation The objective in a problem which displays all of these costs is to determine the number and location of new plants in such a way that the sum of the fixed plant costs plus the fixed route costs plus the variable transportation costs is minimized. This can be expressed as a mixed integer programming problem in the following way.

$$\text{Let } Y_i = \begin{cases} 1 & \text{if plant } i \text{ is selected} \\ 0 & \text{otherwise} \end{cases}$$

$$Z_{ij} = \begin{cases} 1 & \text{if route from plant } i \text{ to warehouse } j \text{ is opened} \\ 0 & \text{otherwise} \end{cases}$$

F_i = fixed overhead cost associated with plant i

R_{ij} = fixed cost of setting up distribution route between plant i and warehouse j

X_{ij} = quantity shipped from plant i to warehouse j

C_{ij} = per-unit variable manufacturing and shipping cost associated with producing in plant i and shipping to warehouse j

The objective function can be written as

$$\text{Minimize } C = \sum_{i=1}^{m} F_i Y_i + \sum_{i=1}^{m} \sum_{j=1}^{n} C_{ij} X_{ij} + \sum_{i=1}^{m} \sum_{j=1}^{n} R_{ij} Z_{ij}$$

The first term represents the sum of all fixed costs for those plants which are selected. The second term is the sum of the variable transportation costs between open plants and warehouses, and the third term is the sum of the fixed costs associated with opening shipping routes between the opened plants and warehouses.

The first set of constraints ensures that the quantity shipped from plant i to all warehouses is less than or equal to the supply available at that plant S_i.

$$\sum_{j=1}^{n} X_{ij} \leq S_i Y_i \qquad \text{for } i = 1, 2, \ldots, m$$

If plant i is not opened, $Y_i = 0$ and $\sum_{j=1}^{n} X_{ij}$ must equal zero.

The second set of constraints ensures that the demand D_j from each warehouse will be met.

$$\sum_{i=1}^{m} X_{ij} \geq D_j \qquad \text{for } j = 1, 2, \ldots, n$$

The third set of constraints relates the value of Z_{ij} to the value of X_{ij}. It is necessary that $X_{ij} > 0$ only if $Z_{ij} = 1$. That is, items can be shipped from plant i to warehouse j only if the route between them is selected to be open. This can be accomplished by the following constraint:

$$X_{ij} \leq U_{ij} Z_{ij} \qquad \text{for all } i \text{ and } j$$

where U_{ij} is a sufficiently large coefficient. For example, U_{ij} could be set to the minimum of (S_i, D_j). Consider formulating this constraint for an example where $i = 1, j = 3$, and

$$S_1 = 15$$

$$D_3 = 10$$

We would have $U_{13} = 10$. If the route from plant 1 to warehouse 3 has been opened, $Z_{13} = 1$ and X_{ij}, according to our last constraint, could take on any value up to 10:

$$X_{13} \leq 10$$

If, on the other hand, the route had not been opened, $Z_{13} = 0$ and

$$X_{13} \leq 0$$

No shipments could have taken place between these two points.

Finally the nonnegativity and integer constraints must be added.

$$X_{ij} \geq 0$$

$$Y_i \leq 1$$

$$Z_{ij} \leq 1$$

where Y_i and Z_{ij} = nonnegative integers for all i and j ////

SUMMARY

When a single facility is to be located and both tangible and intangible costs are present, a simple dimensional analysis approach may be used. If in this case, however, the only relevant costs are tangible and measurable, then a breakeven analysis can be performed.

In multifacility problems the problem is more complex. It must be determined how many new facilities will be added to the existing set and where they should be located. Mathematical programming models have proven to be of great help in solving these kinds of problems.

Even if there are several intangible factors which must be considered in a multifacility problem, it might be helpful at first to set these factors aside and reach a conclusion based solely on measurable costs. Then the intangible factors can be reintroduced and a final decision reached. Often, in fact, these intangible factors will be so strong that they will overshadow the conclusion reached by the quantitative analysis.

In short, facility location requires the skillful blending and weighting of tangible and intangible factors. It is quite unlikely that a mathematical model will ever be developed which is capable of solving a wide range of these problems. As a result, plant location decisions will probably never become routine.

QUESTIONS

Q9-1. What factors would you expect to play a prominent role in the location of the following facilities?

(*a*) State-of-the-art electronic test equipment manufacturing facility
(*b*) High-volume appliance manufacturer
(*c*) Research laboratory for a major pharmaceutical firm
(*d*) Life insurance company headquarters
(*e*) Branch office for a stock brokerage firm
(*f*) Satellite campus for a large urban university

Q9-2. How relevant do you think the following location problems are to the firm's profitability? Explain.

(*a*) Appliance service centers associated with a national appliance manufacturer
(*b*) Supermarket
(*c*) Branch bank
(*d*) Wholesale food broker
(*e*) Steel mill
(*f*) Cable TV company

Q9-3. Why is the problem of locating a single plant different from the problem of locating another plant in an existing plant system?

Q9-4. Should police and fire stations reconsider their location strategy upon a shift in center of population density? Explain.

Q9-5. As a member of the operations research team which investigated the ambulance deployment system for the City of New York, which alternative would you recommend, what interest groups would you involve, and what implementation problems would you foresee?

Q9-6. Under what conditions might dimensional analysis prove useful?

Q9-7. In response to growing demand, a national manufacturing firm has decided to construct an additional plant. Their problem is where it should be located and how large it should be. How would you approach the problem?

Q9-8. Contrast the advantages and disadvantages of large central manufacturing facilities as opposed to those of dispersed decentralized facilities.

Q9-9. What factors do you suppose were considered in comparing sites for the location of Disney World©?

Q9-10. When might breakeven analysis be useful in facilities location?

Q9-11. Many major airports have considered the construction of jetports located some distance from the central city to absorb the forecasted increase in air travel. The tentative choices for these airports are usually in rural or suburban towns. Do you see this as a less painful alternative to central airport expansion? What resistance to this strategy is likely to be encountered?

Q9-12. The optimal locations for missile sites in an ABM system can be determined mathematically. Would you expect public pressure to interfere with this deployment strategy?

PROBLEMS

P9-1. The Shawanagunk Company has three production and three warehouse facilities. Given the following data, find the optimal shipping pattern.

From \ To	A	B	C
1	5	9	2
2	3	7	1
3	2	6	4

Distribution costs

Production capacity		Demand	
Plant 1	20	Warehouse A	15
Plant 2	18	Warehouse B	26
Plant 3	25	Warehouse C	22

P9-2. Solve the following transportation problem.

From \ To	A	B	C	D	
1	3	1	6	4	12
2	5	7	3	6	10
3	7	9	6	5	13
4	4	8	3	5	14
			18	15	49

P9-3. Solve the following transportation problem. Hint: Introduce a dummy warehouse to equate demand with supply. The costs associated with this dummy column will all be zero. Why?

rom \ To				
	7	3	5	
2	4	6	5	
	9	6	7	15
		16	10	

P9-4. Write Prob. P9-3 as a linear programming problem.

P9-5. Solve the following transportation problem. Hint: This is a degenerate problem, since less than $n + m - 1$ assignments will be made by the northwest corner rule. To overcome this, assign a "little bit," ϵ, to a cell so that the remaining empty cells can be evaluated.

From \ To	A	B	C	
	4	6	7	10
2	3	4	7	15
	7	4	9	20
	10	17	18	

P9-6. A large manufacturer currently maintains three factories and three warehouses. The demand from the warehouses, however, currently exceeds the productive capacity of the factories. The figures are given below. Management would like to add a factory to bring productive capacity into line with demand. Under consideration are two locations. From the data presented below determine the optimal strategy.

From \ To	A	B	C
1	7	9	6
2	4	5	11
3	9	4	3
4	8	6	5
5	7	6	2

Transportation costs

Production capacity	
Plant 1	250
Plant 2	130
Plant 3	180

Demand	
Warehouse A	290
Warehouse B	210
Warehouse C	160

Estimated variable per-unit production costs	
Plant 4	$110
Plant 5	$140

Estimated building and equipment costs	
Plant 4	$21,000 per year over the life of the plant
Plant 5	$17,500

P9-7. A large company is in the process of comparing two sites for the location of a new plant. The land cost for site 1 is $1,200,000, while that for site 2 is $600,000. To reflect the importance that this factor should play in the decision process it is given a weight of 5. Building costs are expected to be $2,000,000 for site 1 and $1,000,000 for site 2. This factor also received a weight of 5. Taxes for site 1 will be $100,000, while for site 2 they will be $200,000. This factor received a weight of 2. The labor supply is considered to be adequate in site 1 and acceptable in site 2. It receives a weight of 3. The community attitude for site 1 is ambivalent, but for site 2 it is cooperative. Its weight is 3. Finally, research climate for site 1 is characterized by a few low-quality schools, but for site 2 there are several good research organizations in the vicinity but nothing else. This factor is given a weight of 1. Using Table 12 in the Appendix, determine which alternative is favored.

P9-8. Set up the following mixed-integer programming problem.

To warehouse		
1	2	3
5	6	3
9	4	6
4	3	4

Route set-up costs

To warehouse		
1	2	3
.10	.12	.09
.14	.09	.12
.11	.12	.04

Manufacturing and shipping costs

Fixed overhead costs for plant *i*

i	*Costs*
1	3.5
2	2.7
3	4.1

Warehouse demand	
j	
1	200
2	120
3	140
	460

Supply	
Plant i	
1	145
2	210
3	315
	670

BIBLIOGRAPHY

Babb, E. M., and T. H. Yu: "Analysis of Factors Affecting Least-cost Size of Plant," *Management Science*, vol. 16, no. 10, June 1970, p. B607.

Buffa, Elwood S.: *Modern Production Management*, John Wiley & Sons, Inc., New York, 1969, chap. 10.

Feldman, E. F., F. A. Lehrer, and T. L. Ray: "Warehouse Location under Continuous Economies of Scale," *Management Science*, vol. 12, no. 9, May 1966, pp. 670–684.

Francis, Richard L.: "On the Location of Multiple New Facilities with Respect to Existing Facilities," *Journal of Industrial Engineering*, vol. 15, no. 2, March–April 1964, p. 106.

Garett, J. William: *Production and Operations Management*, Harcourt Brace Jovanovich, Inc., New York, 1968, pp. 225–236.

Karasaka, Gerald J., and David F. Bramhall: *Location Analysis for Manufacturing*, the MIT Press, Cambridge, Mass., 1969.

Wagner, Harvey M.: *Principles of Operations Research*, Prentice-Hall, Inc., Englewood Cliffs, N.J., 1969, chap. 7 and also pp. 451–453.

10

Pollution Control and Its Impact on the Process

INTRODUCTION

In the past, industry has been accused of neglecting its impact on the environment. Quite often the wastes from its transformation processes have overburdened the assimilative capacity of the natural environment. Recently, however, there have been substantial changes which no longer permit these practices. In some cases these restrictions have even jeopardized the viability of the firm.

Because this issue is new and because it often requires a significant restructuring of the transformation process, an entire chapter is devoted to it. First we will try to uncover the reasons why pollution became such a large problem. Then we will look at the different forms of pollution and why the disposal of wastes has gotten out of hand. Following this an attempt is made to quantify waste disposal costs in the hope that a strategy can be achieved to minimize these costs. One of the components of this cost expression, welfare damages, is nearly impossible to quantify, and the usefulness of such an approach is thereby limited. Consequently the policy maker is left with only imperfect strategies. Nevertheless something must be done to prevent unreasonable levels of pollution from occurring and also to control that pollution which already exists. We then turn to the ways in which the public and private sectors of the economy can prevent and control pollution. Next we examine three strategies for establishing formal controls on effluent discharge. Finally, three case studies are presented, one which focuses on a regulated monopoly, another involving a municipal authority, and the last dealing with a private firm.

HOW DID IT GET THIS WAY?

Not many years ago in a certain wooded area located remote from industry and large centers of population, the air was pure, the waters were clean, and there was the noise of only the birds and the wind. As the first few people settled in the area, the water had no problem in assimilating their wastes, the smoke from their fires was easily dissipated in the atmosphere, and the people generated relatively little noise. The population, however, soon began to grow, industry thrived, and the affluence of the community increased steadily. An almost unavoidable concomitant of this developmental pattern was the increase in noise, water, and air pollution.

In the early stages of the town's development, the river was used for swimming, drinking, some waste disposal, and even as an input factor in some chemical and manufacturing processes. The river was truly a multiple-purpose, multiple-use, common-property asset. The first consequence of its heavy use for waste disposal was the curtailment of the use of the river as a drinking-water source. Residents, at that time, had several choices for solving this immediate problem. They could have maintained the status quo and suffered the consequences of dirty water, which might have included increased colds and even more serious afflictions. Secondly, they could have boiled the river water before using it; and thirdly, they could have built a water treatment plant which would have rendered the water drinkable. They chose the treatment plant, which solved the drinking problem; but five years later they had to close the river to all swimming. In addition, several firms which had been using this water in their manufacturing process had to install their own treatment systems to ensure the quality of water which they required.

CAUSATIVE FACTORS: POPULATION AND INCOME

The point of this story was to highlight what have been the two major causative factors of pollution: population density and affluence. As the density of industry and people has increased, the burden placed on the environment has also increased. To compound the problem, as the standard of living became higher, the goods and services demanded per person have increased. We can, therefore, say that the level of pollution is directly related to population density and per capita income. Mathematically we have the following:

$$P = f\ (\text{population, per capita income})$$

where P = pollution level

RESTORATIVE FACTORS: RECYCLING, TECHNOLOGY, AND TREATMENT

This is not nearly the whole story, since we have yet to consider recycling, the consequence of which is to reduce the level of pollution. Consider the familiar controversy associated with soda bottles. The question is whether a no-deposit, no-return strategy should be changed in favor of a recycling strategy. Let us take a closer look at the recycling strategy. Suppose that a large deposit is required on each bottle. Let us also assume that this fee is large enough to ensure the return of most bottles. Once returned, they can be cleaned and reused. The obvious consequence of this is that there is less waste matter which must somehow be collected and disposed of properly or end up as a blight on the countryside. From this example it can be seen that pollution is inversely related to the degree of recycling: as the intensity of recycling increases, the pollution levels decrease.

Technology also acts to ameliorate pollution. Engines which are more efficient and structures which are sturdier but lighter are examples of less waste for the same degree of utility.

Finally, the extent of pollution is inversely related to the intensity of waste treatment. As the percentage of total effluent treated increases, the pollution level decreases. Said another way, higher levels of effluent can be tolerated, provided that it is properly treated by some combination of physical, chemical, and biological action.

Several different kinds of water treatment are commonly employed. These range from mechanical filtration to chemical and biological processing. Mechanical filtration removes just the solids from waste and is the least expensive processing strategy. It, however, provides minimal treatment and is oftentimes inadequate to meet water quality standards. Higher levels of waste treatment are achieved by chemical and biological action.

There are also several ways in which effluents discharged into the air can be treated. The smoke which comes from a power company's stack, for example, can be filtered, scrubbed, or precipitated. The most complete removal of pollutants is achieved by using an electrostatic precipitator. This machine performs its function by establishing an electromagnetic field within which the particulate matter is separated from the smoke.

Noise pollution can be suppressed by the use of silencers at the noise source or by insulating the area affected by the noise. An example can be found in the noise problems associated with aircraft. Newer aircraft models are equipped with silencing equipment, and there is pressure to supply older models with this new equipment. In manufacturing plants it is becoming common to see sound insulation surrounding noisy machines.

POLLUTION FUNCTION

We can now summarize those causative factors which explain the level of pollution. The level of pollution varies directly with the level of population and per capita income, and inversely with the extent of recycling, technology, and waste treatment. Mathematically, this can be expressed in the following way:

$$P = f\left(\text{Pop.}, \begin{array}{c}\text{per}\\ \text{capita}\\ \text{income}\end{array} \frac{1}{\text{recycling}}, \frac{1}{\text{technology}}, \frac{1}{\text{treatment}}\right)$$

POLITICS AND ECONOMICS

Starting in the late 1960s, it became quite apparent that pollution levels had gotten out of hand. Prior to this time, pollution had been a local issue which was confined to the political domain of municipal and sometimes state interest. As the issue rapidly rose on the national priority list, centralization of policy making under the Council on Environmental Affairs and centralization of administration under the Environmental Protection Agency provided the lead to an effective management approach of our common-property resources.

It quickly became apparent that the solution to this problem would be very complex. Although we do know quite a bit about the different kinds of wastes, their chemical composition, and their effect on the environment, we still do not know very much about their long-term effects on the human body. Therefore, the standards which should be established are still subject to controversy. Even when a strategy is selected, the road to its implementation and enforcement is still a rough one. Although economists generally agree upon the economic mechanism necessary to efficiently achieve these objectives, the political factors that must be considered oftentimes dramatically change those strategies. In fact, shortly you will see that the direction that is being taken in pollution control is not the most economical one. Before we examine the economic issues, we take a closer look at the forms of pollution and then at the way in which our environment can and cannot assimilate these wastes.

FORMS OF POLLUTION

WATER POLLUTION

There are several ways in which wastes can be categorized. At the risk of oversimplification but at the same time recognizing the necessity to be brief, we will categorize liquid wastes as being either degradable or nondegradable. The placement of the effluent in either one category or the other will depend upon its behavior in the receiving body of water.

Nondegradable wastes Nondegradable wastes can be diluted and sometimes changed in form, but their weight will not be reduced in the receiving waters. The sources of nondegradable wastes can be found in effluents discharged by industry and runoff from well-fertilized farms. These wastes may remain dissolved in the receiving water or they may settle out. They can be toxic or corrosive, may add color and taste to the water, and may produce odors.

Degradable wastes Degradable wastes can be reduced in quantity by either physical, chemical, or biological action. The most common waste in this category is domestic sewerage.

As long as limited quantities of these wastes are introduced into a body of water, a self-purification process takes place. This process relies on bacteria which are normally present in the body of water. During this decomposition process the bacteria require free oxygen, which is also found in the water. As long as the oxygen is available in sufficient quantity, the decomposition will proceed without producing offensive odors. Oxygen is replenished by being absorbed through the surface of the water and by aeration in rapidly moving streams. If the oxygen is replaced, the assimilative capacity of the river is restored. If, however, the load on the receiving body is too heavy, decomposition takes place anaerobically, or without the use of free oxygen. The consequence of this is that hydrogen sulfide and other gases are produced, a fact which accounts for the offensive odors generated from overtaxed streams.

The depletion of oxygen in a stream has other consequences: most fauna require oxygen to live. Without it, fish life becomes extinct. If the balance of nature is upset and little is done to correct the problem, the eventual consequence is the death of the body of water. One of the most dramatic examples of this is the ecological death of Lake Erie.

Different wastes exert different demands upon the receiving waters. The quantity of oxygen that is required to decompose a particular waste is called its *biochemical oxygen demand* (BOD). The oxygen that is available in a particular body of water, on the other hand, is called its *dissolved oxygen content* (DO). The latter essentially measures how alive the body of water is. Both DO and BOD are crucial factors in any water quality program.

Predicting the time pattern of water quality for a body of water is quite difficult. Consider the case where a single degradable effluent is discharged into a stream. Eventually the water will completely assimilate the wastes, but of importance in water quality management is the progress of the recovery process. The speed with which the degradation of wastes takes place depends upon three factors: the quantity and BOD of the waste, the temperature, and the hydrologic characteristics of the receiving water. The hydrologic conditions are characterized by the flow of the water. The higher the flow, the faster the reaeration of the stream and the quicker the DO is restored to original levels. Temperature also affects the time pattern of degradation, since bacterial action is speeded at higher temperatures.

As the waste is moved downstream, the DO content of the stream starts to decline, reaches a minimum, and finally recovers by hydrologic action to previous levels. This phenomenon is known as the "oxygen sag." To predict the time pattern of this oxygen sag is quite difficult. The problem is that the water levels, water flows, and temperature vary considerably over time for most bodies of water.

Compounding the problem is the fact that several different kinds of wastes are usually discharged into a stream at several different locations. Some of these wastes have simple additive effects. That is, they do not interact with one another. Other wastes have counteracting effects. The combination of X with Y results in a reduction of the damage that would occur if X or Y were introduced to the stream singly. Finally, we have the case where the introduction of two wastes results in an escalation of the damage that would have occurred had the two not been introduced together.

We have already uncovered several factors which make the establishment and subsequent management of a water quality control program quite complex. These include: the setting of standards, some of which are still subject to controversy, and the ability to predict the water quality, given a continual variation in the quantity and kinds of effluents.

Once the decision is made as to the waste load which the water can sustain, some mechanism must be established to accomplish this end at the lowest possible cost. Perhaps the most common mechanism is regulation. Here is where the need for economic analysis becomes a necessity, for regulation is but one of the several alternatives which can be employed to accomplish these objectives, and each has a different cost of administration, cost to the firm, and consequently a different cost to society.

AIR POLLUTION

Unlike water, air is not self-purifying. As a consequence, the only way that air can be cleansed is by cleaner air pushing the dirtier away or by particulate matter settling out from the air and contributing to ground pollution.

A major problem with air pollution is that several of the sources are mobile and therefore difficult to control. The exhaust from cars, trucks, buses, and airplanes is an example of this. To obtain effective control over the discharge levels of these forms of transport, it is necessary to establish and enforce federal standards. Nonmobile sources of pollution include electric power plants, municipal and residential rubbish burning, industrial plants, and home heating systems. Each makes some contribution to pollution; however, the chemical and particulate composition of their discharges is quite different. A breakdown is shown in Table 10-1.

Table 10-1 Air pollution emission in the United States†

			Sources—millions of tons per year				
Pollutant	*Million tons per year*	*% of total*	*Power plants*	*Industry*	*Motor vehicles*	*Space heating*	*Refuse disposal*
Carbon monoxide	72	50.7	1	2	66	2	1
Sulfur oxides	26	18.3	12	9	1	3	1
Nitrogen oxides	13	9.2	3	2	6	1	1
Hydrocarbons	19	13.4	1	4	12	1	1
Particulates	12	8.4	3	6	1	1	1
Total	142	100.0	20	23	86	8	5
Percent of total			14.1	16.2	60.6	5.6	3.5

† "The Sources of Air Pollution and Their Control," U.S. Department of Health, Education, and Welfare, Washington, D.C., 1967.

NOISE POLLUTION

Little is really known about noise pollution. We need to know more about the harmful physiological and psychological consequences which prolonged exposure can cause. The initial arena for focus on this problem was airport-related noises. With the increase in air traffic and the higher levels of noise associated with jet engines, airport administrators have had to face an increasingly irate local citizenry. Some progress has been made by the forming of coalitions of interested parties, including, of course, citizens from these local communities. We will return to a case study of one such problem at the end of this chapter.

WHY HAS WASTE DISPOSAL GOTTEN OUT OF HAND?

The production process involves the combination and subsequent transformation of several resources. These include men, machines, material, and even air and water. To some extent one can be substituted for the other. A machined product, for example, can be processed mostly by skilled labor with the use of manually operated drills, lathes, and grinders, or it could be processed completely by numerically controlled machines. In the first case, the process is labor-intensive; in the second, it is capital-intensive. In another example, power can be produced by relatively inexpensive high-sulfur coal, discharging large quantities of both sulfur dioxide and particulate matter into the environment, or it can be generated by expensive fuel oil, with much lower levels of discharge.

These examples point out that there are usually many ways in which resources can be combined to produce a product. The particular combination which is finally chosen depends upon their relative prices. For example, if skilled labor is much more costly than machine time, the process will be capital-intensive. If, on the other hand, machine time is more costly, then it will be labor-intensive. As the price of a resource falls relative to other substitute resources, its use will increase.

Now reconsider the choice facing the power company. On the one hand, they can use relatively inexpensive coal but discharge undesirable effluents into the air; or, on the other hand, they can use expensive fuel oil and discharge quite acceptable levels of effluent. If we assume that "using air" or discharging effluent into the air is costless, any cost-minimizing (in this case) or profit-maximizing (as in the case of a nonregulated firm) firm would elect to purchase coal. Only when a charge is imposed for using the air might the firm shift away from its present strategy of using the inexpensive coal.

From this we can conclude that resources which have little or no cost to the firm will be used much more extensively than if they involved a cost. The problem, then, has been in the inability of the market to impose a cost on these resources. As a result, the total cost of producing goods which require air and water as input resources has been less than it should be. Since these goods are subsequently sold at a lower price, it seems reasonable to hypothesize that we sell relatively more goods which require air and water as input resources than we would sell if they carried their full cost. We have clearly encouraged pollution!

Today it is transparently clear that there is a positive cost when air and water are used as input resources. That this is true can be seen by tracing the effects of the discharge from a power plant. The smoke contains, among other things, sulfur oxides and particulate matter. Both are harmful to health. Particulate matter increases cleansing bills, and in general the haze that the discharge causes is plainly uncomfortable and unsightly. When the power company uses inexpensive fuel, it is not incurring these external costs. The cost to the power company of using the air is virtually zero. However, the other users of the same air must incur some extra costs. A major problem, then, is that the users of the environmental resources are not necessarily the ones who suffer the damages. Moreover, the ones who suffer the damages are generally large in number, dispersed over a geographic area, and unlikely to take collective action. As a consequence, the use of these resources remains free to the discharger until such time as the cost of the damages becomes so great that political and economic action is eventually taken and a cost is imposed on their use.

WASTE DISPOSAL COSTS

A WASTE DISPOSAL MODEL

A conclusion which we can reach is that waste disposal costs affect several people and the costs which are borne by the discharger are not the only costs that should be taken into consideration—there exist a set of external costs. To be a little more explicit, we can define waste disposal costs for any given level of waste disposal as the sum of pollution prevention costs and pollution costs.[1]

$$\begin{matrix}\text{Waste disposal}\\ \text{costs}\end{matrix} = \begin{matrix}\text{pollution}\\ \text{prevention}\\ \text{costs}\end{matrix} + \begin{matrix}\text{pollution}\\ \text{costs}\end{matrix}$$

Pollution prevention costs are those expenditures incurred by public and private enterprises to prevent some of the damaging effects which can be attributed to wastes. Examples include the expenditure of funds for sewage treatment plants by municipal authorities, the processing of effluents by manufacturing firms to remove harmful chemicals, and the expenditures by individuals to purchase and maintain pollution control devices on their automobiles. In general, then, pollution prevention costs are expenditures incurred to prevent the development of a pollution problem.

Pollution costs are those costs incurred after the pollutants have been released into the environment. They represent the costs which are borne because of a pollution problem. As you will shortly discover, these costs can be quite difficult to measure. Pollution costs can be broken down into three categories. These include public and private expenditures to avoid pollution damage, and the welfare damage that is suffered rather than prevented.

---

[1] J. H. Dales, *Pollution, Property and Prices,* University of Toronto Press, 1968, p. 15.

$$\begin{array}{c}\text{Pollution}\\ \text{cost}\end{array} = \begin{array}{c}\text{Public ex-}\\ \text{penditure to}\\ \text{avoid pollu-}\\ \text{tion damage}\end{array} + \begin{array}{c}\text{private ex-}\\ \text{penditure to}\\ \text{avoid pollu-}\\ \text{tion damage}\end{array} + \begin{array}{c}\text{welfare}\\ \text{damage}\\ \text{of pol-}\\ \text{lution}\end{array}$$

Public expenditures to avoid pollution damage refer to those expenditures made to avoid the consequence of using polluted air or water. A water-purifying facility would fit into this category. A sewage treatment plant, however, would not, since its purpose is to prevent pollution. We would classify it as a pollution prevention cost.

All private expenditures to avoid pollution damage are contained in the second category. These include expenditures for air and water purification systems that industry must install to meet process standards. Also included are the extra cleaning bills, house painting, medical bills, travel expenditures for vacationing in unpolluted areas, and the cost of home water purification systems which those who live in polluted areas must incur.

The last and most baffling category of pollution costs is the welfare damage that is suffered rather than prevented. Frequently pollution prevents the enjoyment of certain leisure activities. Such things as swimming, water skiing, and fishing represent pleasure that must be forgone when a body of water becomes polluted. Exactly how the dollar value of this loss should be calculated is a very difficult if not altogether impossible question to answer. Usually these activities are not lost altogether. Vacationers and sportsmen, however, must travel further to enjoy unpolluted waters.

Another example of welfare damage that is suffered rather than prevented is found whenever a family decides to move out of an urban setting into a less polluted environment. Perhaps the salary of the wage earner is less as a result, or, if he commutes to work, his expenses may be greater. Firms also may move out of the city for much the same reason.

A final example of welfare damages is the mental and physical discomfort associated with pollution. Dirty air, unpleasant odors, unsightly streams, and noisy communities all involve some discomfort which many people would pay to alleviate. These aspects of pollution are not necessarily serious health hazards nor do they result in expenditures for dry cleaning, house painting, and so on. Those consequences of pollution we have already considered. These are just the nuisances that do, however, detract from the quality of life.

We can now restate the waste disposal costs in the following way:

$$W = P + U + V + D \tag{10-1}$$

where W = waste disposal cost
P = pollution prevention costs
U = public expenditures to avoid pollution damage
V = private expenditures to avoid pollution damage
D = welfare damage

Now we turn to the problems which are encountered in the design of a pollution control program.

THE MINIMIZATION OF WASTE DISPOSAL COSTS

The central issue in pollution control is to determine the proper strategy for disposing of society's wastes in such a way that the waste disposal costs are minimized. What this implies is that a correct balance must be struck among pollution prevention costs, public expenditures to avoid pollution damage, private expenditures to avoid pollution damage, and welfare damage. Consider an example where a firm is about to dump 300 pounds of effluent into a stream. If this waste is dumped untreated, welfare damage estimated to be $100 will occur. There are, however, two other possible outcomes which have been evaluated. If the firm treats the waste, the pollution prevention costs incurred will be $65, and it is expected that no welfare damage will result. Finally, the firm can dump its effluent into the stream and the local municipality can treat the water for drinking purposes at a cost of $50 while $25 of welfare damage will remain. The waste disposal cost for each strategy can therefore be computed in the following way:

	P	+	U	+	V	+	D	=	W	
Strategy 1	0	+	0	+	0	+	100	=	100	
Strategy 2	65	+	0	+	0	+	0	=	65	←
Strategy 3	0	+	50	+	0	+	25	=	75	

It can be concluded that the strategy which involves an expenditure of on pollution prevention is the most economical.

Consider another problem, one in which only two strategies are relevant. The first solution is to let the effluent go untreated and suffer a welfare cost of $200, while the second is to treat the effluent before it is discharged into the stream, at a cost of $350. In this example, it is clear that the most economical solution is to let the effluent go untreated. It makes little sense to spend $350 in order to avoid $200 worth of damage. For some, the thought of permitting any welfare damage to occur is unacceptable. But since there is a cost which can theoretically be assigned to the welfare damage of waste practices, it is quite reasonable to expect that certain levels of welfare damage are not worth avoiding.

We can therefore conclude that the objective in waste disposal is to minimize these costs for a given level of wastes. Costs include the welfare damages which are suffered, and it is inevitable that some welfare damage will always be incurred; the question is, How much? To answer this is very difficult and depends on our ability to measure these costs. Once measured or estimated, our theoretical construct would lead us to evaluate each of the available strategies and choose whichever one minimizes waste disposal costs. Now we take a closer look at something I am sure has been bothering you: How do you measure welfare damages?

WELFARE DAMAGES

It is nearly impossible to determine the dollar value of welfare damage associated with a level of waste disposal. Consider the prevention of leisure activities, such as swimming, water skiing, and fishing. It would be quite difficult to determine the dollar value of damage that is suffered when a lake can no longer be used for these purposes. One could conceivably ask people how much they would be willing to pay in order to avoid this problem, but even that might not work. People may understate costs if they feel they have to pay, or overstate them if they think the state will pay.

Complicating the problem is usually the proximity of substitutes. In the case of the lake, there may be another one just like it but quite suitable for these activities just 90 miles away. The relevant loss which is incurred is much smaller than the complete loss of the polluted lake to vacationers. It is the difference between the cost of travel to the polluted lake and that of travel to the unpolluted lake. There may be some, however, who cannot afford to travel this extra distance. For them the loss of this leisure activity is complete.

Voting as a means of resolving problems which are laden with welfare considerations is also not without problems. Consider a pollution problem which can be solved if the municipality builds a sewage treatment plant at a cost of $100. For purposes of simplicity, assume that there are only three voters and that each is affected differently by the problem. The assessment for the sewage plant, however, would be split equally among the three residents at $33.33 each. Voter 1 carefully considers the amount of money that he would be willing to pay in order to avoid the welfare damage which he suffers. He concludes that he would be willing to pay up to, but not more than, $20. As a result, he decides not to vote for the sewage plant. Voter 2 also assesses his damage as $20, but Voter 3 feels he would pay $80. The only favorable vote is from Voter 3, and the sewage treatment proposal is defeated. This defeat takes place despite the fact that collectively the voters would have to pay only $100 in assessments to avoid a problem which is costing them $120 in welfare damages.

As you can see, it is nearly impossible to determine the welfare damage component of waste disposal costs. Consequently we do not have a basis for a clear-cut quantitative comparison of alternative strategies to meet a pollution problem. This does not mean that our conceptual framework which has been built around the expression for waste disposal costs in Eq. 10-1 is valueless. It does provide a valuable framework within which the problem can be analyzed. It should be clear, however, that no "optimal" solution to the problem exists.

The pollution control strategy which is finally chosen should be the one which comes closest to minimizing these costs. A strategy for a particular watershed, for example, would include a set of water quality standards. It is to be hoped that, when these water quality standards are imposed on the users of the watershed, the waste disposal costs, including welfare damages, would be approximately minimized.

POLLUTION PREVENTION ALTERNATIVES

The issue which must be considered next, and an issue which is much more tangible than the initial setting of the standards, is the identification of the alternative pollution prevention techniques which can be employed to achieve these standards. These alternatives can be divided into two categories. The first includes those alternatives open to the management agency of the resource, such as a watershed. The second, those alternatives which the firm can employ to manage the quality of its effluent.

MANAGING AGENCY

A management agency whose charge it is to achieve and maintain a certain quality of water can employ some combination of waste treatment plants, the addition of dilution water whenever river flows are low, and reaeration of streams to add oxygen by bubbling air through the water. The waste treatment plant is an example of reducing wastes after they are generated, whereas the dilution and reaeration techniques represent methods for increasing the assimilative capacity of the water.

Since the pollution level of a river is rarely stable, the managing agency can use these alternatives in some economically inspired combination to maintain the standards which have been established. For example, rather than building a waste treatment plant to accommodate the highest concentration of waste, which may occur only a small percent of the time, it might be more beneficial to build a dam and utilize the water for dilution purposes during peak periods of effluent flow. The complex nature of this aspect of the water quality management task should now be apparent to you. In summary, the essence of it is this: A strategy must be selected from among several alternatives to achieve a set of water quality standards at minimum cost. Because of the variable nature of the river flow, of the effluent discharges, and of the water temperature, the response made today may not include the same set of alternatives used tomorrow.

FIRM

A second set of alternatives for achieving a specified effluent quality are under the control of the firm. They can be considered as either methods for reducing waste generation or methods for reducing wastes *after* generation. First we consider the alternatives for reducing waste generation.

Frequently the type of raw material employed in the transformation process affects the nature of the waste. An example which was mentioned earlier is the case of power production. If coal which is high in sulfur content is used one level of effluent is incurred, whereas if oil is used the effluent is much cleaner.

Another alternative is to make a change in the transformation process. This might involve the reduction of scrap, the substitution of one chemical process for another, or the use of air-cooling systems rather than water cooling. The steel industry, for example, for years had used sulfuric acid for cleaning steel. A major problem, however, was the quantity of effluent which this process generated. After switching to a hydrochloric acid solution, the effluent generated was substantially reduced. Another example can be found in airport management. To ameliorate the noise pollution found at an airport, several steps can be taken. These might include the use of certain runways at night to keep the noise over populated areas to the lowest possible levels, the restriction of power runups, and the maintenance of high altitudes before landing and of steep ascents after takeoff.

Now we take a look at methods for reducing wastes after they have been generated.

First, it might be possible to recover valuable materials which might otherwise be discarded. A good example of this can be found in the plating industry. After a part has been electrochemically plated, the plating solution must be rinsed off in a water bath. The disposal of the water rinse which contains toxic chemicals presents a serious pollution problem. One solution is to evaporate the rinse water and reuse the residual in the plating bath.

Another strategy for removing wastes after their generation is to produce salable by-products. These might include either consumption or intermediate goods. Examples include the by-products of monosodium glutamate and potash from beet sugar production, utilizing the wastes from apple canning for vinegar production, and the use of cottage cheese whey for protein food supplements.

These last two strategies—materials recovery and by-product production—are seldom profitable enough to induce the firm to undertake them in the absence of either some penalty for discharging their effluent or a sizable public conscience.

A third, and perhaps more common, strategy involves the use of waste treatment. This can occur either by the firm's treatment of their own wastes or by their use of municipal or shared waste treatment facilities. Whether these facilities are private, shared, or municipal will, of course, depend upon the costs associated with each alternative. Since, however, there are economies of scale in waste treatment, shared or municipal facilities can be an economical solution for many waste problems.

WASTE TREATMENT: PRIMARY AND SECONDARY

There are two basic ways of treating wastes. These include primary and secondary processes. In a primary treatment process, the solid matter in the effluent is filtered and the particulate matter is settled out. It represents a very minimal and usually wholly inadequate means of treating wastes. Secondary treatment, on the other hand, combines the primary process with bacteriological action. The resultant water quality is substantially better than that from the simple primary process.

The major characteristics of a primary treatment facility are shown in

Fig. 10-1. The effluent is first passed through a wire-mesh screen. Here the large, floating objects are removed. After screening, the sewage passes into a grit chamber. In this chamber sand, grit, cinders, and small stones settle to the bottom. The sewage that remains still contains small particles that are suspended in the solution. These are removed by treatment in a sedimentation tank. This is accomplished by reducing the effluent flow through the tank so that the particles sink to the bottom. The solid mass that accumulates on the bottom is called *raw sludge;* it must be removed, together with the sediment collected from the screening process and the grit chamber. After completion of the sedimentation process the effluent is treated with chlorine to remove harmful bacteria and to reduce objectionable odors. It is then discharged into a stream or river.

The water which is discharged from a secondary treatment facility is of higher quality. Up to 90 percent of the organic matter in sewage is removed by this process. Before the effluent is subjected to secondary treatment, however, it must first pass through primary treatment. After the effluent leaves the sedimentation tank it can be processed in one of two ways. In the first, it is sprayed over a bed of stones. This bed is called a *trickling filter* and is from 3 to 10 feet deep. Bacteria gather and multiply on these stones, and as the sewage is passed through, these bacteria consume most of the organic matter. The final step in the secondary treatment phase is the addition of chlorine to the water before it is finally discharged into a river or stream.

Another type of secondary treatment process utilizes activated sludge. This process, illustrated in Fig. 10-2, speeds the bacteriological action by intro-

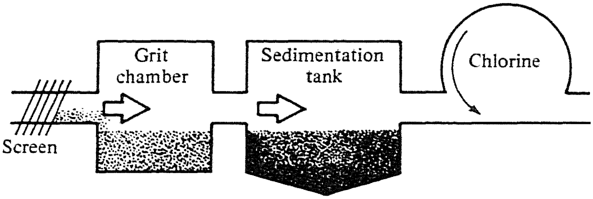


Fig. 10-1 Primary waste treatment.

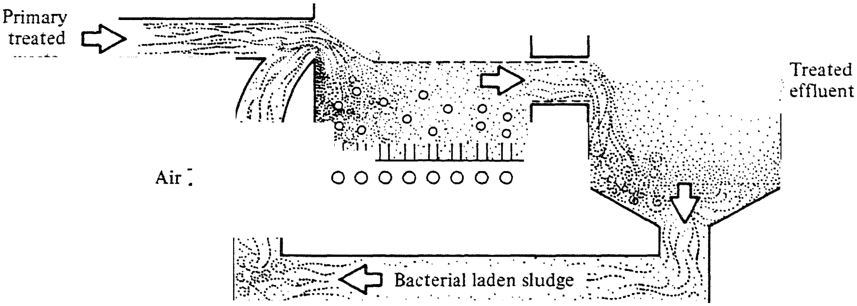


Fig. 10-2 Secondary waste treatment.

ducing air and sludge heavily laden with bacteria into close contact with the sewage.

THREE STRATEGIES FOR CONTROLLING LEVELS OF EFFLUENT DISCHARGE

Once the tolerable level of effluents has been determined by the managing agency, the problem is how to achieve these levels at minimum cost. Three strategies have been suggested. They include control by regulation, control by subsidy, and control by charges. The economic consequences of each strategy are different; therefore, it does make a difference which one is finally chosen. First we look at control by regulation.

REGULATION

Regulation involves such mechanisms as licenses, permits, compulsory standards, zoning, and registration. It can be implemented in two ways. These include across-the-board and point-by-point regulation. In across-the-board schemes, all polluters are treated as if they had an equal share in the damage process. For example, if a regional watershed authority desired to reduce pollution by 10 percent, under an across-the-board scheme they would require all polluters to cut back exactly 10 percent. That this is grossly inefficient and unfair can be seen from a few illustrations. It is inefficient since it is unlikely that the marginal cost of reducing pollution is the same for all polluters. It might very well be that for two firms located side by side on the same river, the cost incurred by one of the firms is double that of the other. The cost to society would be much less if the other firm alone reduced its effluent to meet the new standards.

Across-the-board regulation can also be quite unfair. Consider the case where some firms have, in the past, spent a considerable sum on pollution control. When the across-the-board reduction was announced, it would not be unreasonable to expect that these firms would feel that those who have done nothing in the past should be required to do much more than those who have been cutting down their pollution.

In general, then, across-the-board regulation is neither efficient nor equitable. Now we turn to another form of regulation, point by point, to find that its shortcoming is administrative costs.

In point-by-point regulation, the damages inflicted by each polluter become the basis for his imposed constraints. For this strategy the pollution generated by each firm must be examined and its downstream or downwind damage evaluated. At the same time that his effluents are examined, the costs to him for cutting back these effluents would be estimated. The regulatory body would then know the amount of damage and the cost of cutting back on these effluents for each source. A strategy could then be determined which could minimize the waste disposal costs for a given standard of air or water quality. In theory, then, point-by-point regulation sounds feasible. However, when administrative costs are considered, it becomes unreasonably costly. The size of the bureaucracy that would be needed to determine, administer, and monitor the thousands of regulations would be overwhelming.

SUBSIDY

Subsidies are viewed quite favorably by local municipalities and industries because local taxes do not have to be raised, nor do prices. Typical of a subsidy is the 75 percent that the federal government will give to the municipality toward the construction costs of waste treatment plants. Other kinds of subsidies include the accelerated depreciation that is allowed on all pollution control equipment, the abatement of local property taxes on pollution control equipment, and payments made to the polluter for decreases in the quantity of effluent discharged.

Suppose an equal sum was paid to each polluter to reduce his wastes by 10 percent. This would be an example of across-the-board subsidization. Would this be an efficient means of reducing pollution? No! As we discussed earlier, the cost of reducing effluent levels varies among all polluters. For some the subsidy would be more than the cost of cutting back on their effluent, and they would make money. For others, however, the subsidy might be far less than the cost.

A major problem associated with this kind of subsidy, however, is that those firms whose waste reduction costs are much lower than the subsidy would be inclined to move into the region covered by the subsidy and make money on their wastes. If this seems farfetched, how about farmers who receive subsidies for not growing crops?

An even more efficient approach would be to identify those polluters who could reduce their effluent most efficiently, and to subsidize only those and by an amount equal to the cost of their efforts. The problems inherent in this point-by-point subsidy strategy are precisely those which accompany point-by-point regulation. These include high costs of administration, a sizable bureaucracy, and a complex and costly investigation, negotiation, and monitoring process associated with the polluters chosen to participate in the abatement process.

CHARGES

Finally we look at the most *economically* feasible alternative: pollution charges. In these schemes a charge is levied for an equivalent amount of waste discharged into the watershed. A table would have to be compiled so that by means of it the particular effluent discharged at a plant could be converted into its equivalent amount of waste. Once the charge had been established, those firms whose treatment costs exceeded the charge would continue to discharge into the watershed, whereas those firms whose treatment costs were less would refrain from doing so and treat their effluent.

A point-by-point charge makes the most economic sense. Here the effluent charge is based on the downstream damages that the effluent will inflict. When this charge is imposed, the firm will take steps to reduce its discharge by some

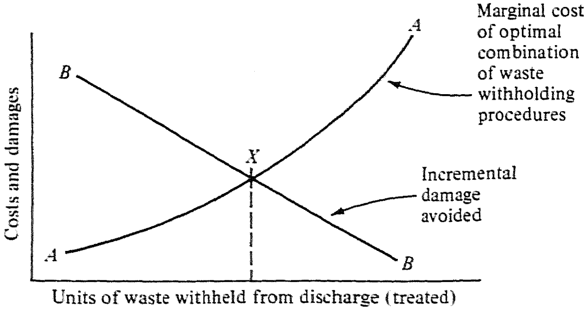


Fig. 10-3 Determination of waste-withholding level.

optimum combination of product or process adjustments, waste treatment, etc., until the marginal cost of another unit of abatement equals the damage.

This concept can be seen graphically in Fig. 10-3. Curve *AA* represents the marginal cost of the optimal combination of waste-withholding measures. It is the least costly way in which the firm can reduce the next unit of effluent. The line *BB* represents the damages avoided when a certain quantity of waste is withheld from discharge. It can be seen that the first few units of waste which are withheld or properly treated result in the avoidance of large downstream damages. As the quantity of wastes withheld or treated increases, these avoided damages decrease as the assimilative capacity of the water performs its natural purifying function.

The firm faced with a cost function *AA* and a charge imposed which is equivalent to the incremental damage avoidance function *BB* will treat wastes to a level *X*. At that point the sum of the treatment costs and effluent charges would be minimized.

Specifying the incremental damage avoided at various levels of treatment has associated with it all the welfare problems that are so difficult to quantify. From a practical point of view, these charges must nonetheless be determined in the best way possible. One way to accomplish this would be to undertake a period of experimentation where charges would be escalated until the proper quality of water or air was achieved.

WHO PAYS?

Under all of these schemes—regulation, subsidy, and charges—the cost of cleaning up the environment is eventually shared among the taxpaying and consuming public. However, the way in which these costs are shared depends upon the scheme which is chosen. Under across-the-board regulation, the consumers of the particular products bear the brunt of waste disposal costs; but since the disposal costs are higher than they would be under another scheme, the cost to society for these products is higher than they could have been. Under a subsidy scheme the taxpayer pays the cost of pollution, and if he does not consume the particular products affected, he is subsidizing those who do. In addition, since taxes are not neutral, the cost of cleaning up the environment is not shared equitably by society.

Under point-by-point charges, the consumers of the particular products bear the full costs of production. The product will be priced to include its traditional manufacturing costs as well as the cost of the common-property resources which it uses.

The revenue which is generated by a pollution charges scheme can be used for administrative purposes as well as for centrally treating those wastes which industry finds it cheaper to dump than to treat. This program can therefore eventually be self-supportive.

Now we turn to focus strictly on the micro phase and examine the way in which the firm responds to these issues. To accomplish this we look at three brief case studies: one in water pollution, the other in air pollution, and finally one in noise pollution.

Case study 1: Metro Power and Light Company Metro Power and Light Company supplies 20,000 customers with electric power. In the past, the company had purchased fuel for their boiler at the lowest cost possible. The sulfur oxides and fly ash which were discharged into the air from this low-grade coal were, by today's standards, quite unacceptable. New standards, however, which were imposed two years ago led them to clean up their effluents. To meet a 3 percent sulfur and .8 percent fly ash standard, they had to change to Number 6 residual fuel oil and to purchase a cyclone collector. The fuel itself barely met new sulfur standards, while the cyclone collector was needed to reduce fly ash content to acceptable levels. The net annual cost of these pollution control steps increased operating costs by $160,000, or an average of $8 per customer per year.

Since Metro Power is a regulated monopoly, they should be able to eventually pass additional expenses on to their customers. This is not, however, automatic. First they must apply to the state utility commission for a rate review. There is never a guarantee that part or all of the rate increase will be approved. This time, however, the rate increase is approved. Therefore, the ones who are absorbing the cost of pollution prevention were the ones who are benefiting from the service—just as it should be.

For Metro Power, however, the problems are far from over. New standards will shortly apply. The new sulfur level has been reduced to 1 percent and fly ash to .2 percent. There are some promising new processes which as yet are not operational. Until they become available, however, the best strategy would be to burn Number 6 residual fuel oil together with an additive and then process the stack effluent through a 95 percent efficient electrostatic precipitator, which will remove enough fly ash to meet the new standard. This process will add another $150,000 per year to the expenditures already being incurred to meet present standards. The per-customer increase, if approved, will be $7.50. In total, then, the customers will be paying an average of $15.50 per year in pollution prevention costs. Presumably the increase in pollution prevention costs of some $310,000 over old methods is less than the welfare damages that would have been suffered had nothing been done.[2] Otherwise more money is being channeled to pollution prevention than is socially warranted.

Case study 2: Logan International Airport Logan International Airport is one of the major airports in the United States. A consequence of this is that it is a noisy place. For the communities surrounding the airport, the situation is becoming less and less tolerable.

The NEF 30 noise contour within which noise levels are considered to be annoying encompassed 25 square miles and 177,000 people in 1967. By 1975, this contour will encompass 80 square miles and 556,000 people. The area least compatible with aircraft noise is considered to be within the NEF 40 noise contour. It will increase from 5½ to 13 square miles and the number of people affected will increase from 30,000 to 76,000 over this same period.

The local citizenry have not taken this sitting down. A group composed of several interested parties has been formed to investigate problems and suggest solutions. It includes representatives from the Massachusetts Port Authority (the airport administration), the Airline Pilots Association, the Federal Aviation Administration, the Massachusetts Aeronautics Commission, and several surrounding communities.

The airport has succeeded in implementing several operational procedures for reducing noise levels. These include such strategies as steep ascents and descents of aircraft so as to maintain high altitudes for as long a time as possible, preferential use of runways to limit the use of those which impose high annoyance levels on surrounding communities, restrictions on power runups, maximum use of over-water flight paths, and procedural turns on takeoff to avoid the most populated areas.

Operational procedures are not the only means by which annoyance can be reduced. Nonoperational procedures, including such strategies as quieter aircraft engines, and effective management of land use in the vicinity of the airfield have been considered.

The cost of quiet aircraft is quite high. The problem is that the quieter the aircraft, the more inefficient and costly it is to operate. Furthermore, it is impossible for Logan Airport, by itself, to require silencers on all aircraft. Such a move must be initiated on the federal level.

A more drastic nonoperational strategy is to redevelop the surrounding region. This might include purchasing homes, industries, and schools, arranging for large-scale relocation, or even insulating structures where the annoyance level is high.

In an appropriately zoned airport region, the mix of residences, industries, hospitals, playgrounds, and so on would be skillfully designed so that annoyance would be kept to tolerable levels. Since the occupants of each different kind of structure have a different annoyance threshold, there exist certain layouts which are much better than others.

---

[2] We assume that no changes are made in private and public expenditures to prevent pollution damage.

A well-conceived solution to this noise problem requires the consideration of both operational and nonoperational strategies. No solution, however, will completely eliminate the noise; some level of noise is inevitable. Prevention measures should be undertaken until the marginal cost of noise prevention equals the marginal welfare damage avoided. At that point an economic solution to the problem has been achieved.

At present, however, there are no standards which guide noise abatement efforts at Logan. Progress is slow, and little has been done with the exception of the operational procedures.

Recently the airport has announced a major expansion program. Two new runways will be added and another two will be expanded. The Massachusetts Port Authority maintains that the expansion is necessary to meet future demand. Without it, airport capacity will be exceeded within the next four years. As you would expect, the citizens from the surrounding communities are violently opposed to what they see as the further deterioration of their environment.

Case study 3: Good Milk Company Good Milk Company, one of the largest milk companies in New England, processes and sells milk and cheese products. A by-product of the cheese process is a waste called *whey.* To make 10 pounds of cheese it takes 100 pounds of milk, while the remaining 90 pounds is whey. In the past the whey has been dumped into convenient rivers and streams. Eventually it decomposes in the water; however, during the process it demands large quantities of oxygen—that is, the BOD requirement for whey is high.

Recent state legislation has established effluent standards which essentially prohibit their plants from dumping untreated whey into streams. These plants are suddenly faced with the prospect of choosing between several alternatives.

The first strategy is to construct their own whey treatment plant. Operating and capital costs would run about 30 cents per hundred pounds of raw whey for the average plant. This is the most expensive alternative.

A better strategy would be to connect into a municipal system. In fact, one of their cottage cheese plants is now undergoing negotiations with local officials. The town of 1,400 people, of which 100 are employed by Good, is also in violation of state pollution standards. They currently have no waste treatment facility. Good has suggested that they both cooperate on a joint venture. They have guaranteed the local townspeople that the project will not cost them any more than if they had gone into it alone. The costs of whey treatment under this strategy will be approximately 14 cents per hundred pounds of whey.

Another strategy is being considered for one of their cottage cheese plants in Vermont. The majority of the 16 cheese plants in Vermont may enter a cooperative for processing whey into a salable powder. Such a powder is rich in vitamins and minerals and has been used as an ingredient in foods. At present, however, only 20 percent of the whey produced across the country is sold for human consumption in this form.

This conversion process is quite expensive, and it is therefore necessary to have several cheese factories over which to spread fixed costs. In addition, marketing the powder will be a task of considerable magnitude, and this can also benefit from a centralized organization.

Given that the cheese processors pay the central whey processor approximately 10 cents per hundred pounds of whey, the process, including all transportation, selling, and administrative costs, will break even. The hope is, however, that as new markets are found for the whey powder, the profitability of the process will increase. Recently, for example, bakers have started to use it in the bread-baking process.

Time is running out; something must be done. Although few water pollution cases have been brought to New England courts before 1971, the increase in number now scheduled indicates that these states mean business.

SUMMARY

In this chapter we considered the reasons why the problem of pollution grew to such significance. Essentially we placed the blame on the failure of the marketplace to impose costs on the use of air and water. To solve the problem it was suggested that one of three methods be employed. These include regulation, subsidy, and charges. Substantial progress has been made in employing two of these methods: regulation and subsidy.

Already the Environmental Protection Agency, backed by the legal system of the United States, has been given broad regulatory power of pollution-related activities. Since other applications of the regulatory concept have never met the aspirations of most economists, we must be especially careful of how this new program is designed and administered.

Needless to say, many firms are faced with substantial problems. Often their processes must be redesigned, their sources of supply changed, and the way in which they dispose of their effluent changed. It will be several years before the transitionary phase is over and environmental standards become common, accepted, and uncontroversial.

QUESTIONS

Q10-1. On what does the level of pollution depend?

Q10-2. What is the difference between degradable and nondegradable wastes?

Q10-3. Why is it difficult to predict the oxygen sag of a stream?

Q10-4. Should the quality of the receiving air or water influence the level and concentration of waste which is allowed to be discharged into it?

Q10-5. What is the basic difference between the assimilative capacity of air and that of water?

Q10-6. Would you attribute the deterioration of our environment to callous entrepreneurs or the failure of the market mechanism to place a cost on the use of air and water?

Q10-7. How can a price be imposed on our natural resources so they will not be used as a free good?

Q10-8. Write an expression for noise "disposal" cost. Explain the significance of each term in the expression.

Q10-9. Why are welfare damages so hard to determine?

Q10-10. In the minimization of noise disposal costs, is it desirable to reduce welfare costs to zero?

Q10-11. Describe the primary and secondary waste treatment processes.

Q10-12. Why is across-the-board regulation so inefficient?

Q10-13. Why is the charges scheme the most economically attractive?

Q10-14. In the Logan Airport case, assume that you are working for the Massachusetts Port Authority. What strategy would you follow?

Q10-15. If you were to speak with an executive from Metro Power, he would certainly not agree that the higher cost of low-sulfur fuel can be passed on to the customer easily. He would be quick to point out that the power company can raise their rates only upon approval by the state Public Utility Commission. In fact, several years may elapse before a rate increase is approved. Nonetheless the power company is entitled over the long run to not only recover expenses, but also to earn a reasonable return on its investment.

With this new information, would you expect any difficulties to emerge when the director of pollution control faces the board of directors with his investment requests?

Q10-16. At the same time that the Vermont whey cooperative venture is showing some promising signs, the state pollution authorities are pressing for some immediate response. What should be done?

BIBLIOGRAPHY

Dales, J. H.: *Pollution, Property, and Prices,* University of Toronto Press, 1968.

Goldman, Marshall I.: *Critical Issues in Controlling Pollution,* Prentice-Hall, Inc., Englewood Cliffs, N.J., 1972.

———, *Controlling Pollution,* Prentice-Hall, Inc., Englewood Cliffs, N.J., 1967.

part two

Operation and Control of the System

11

Forecasting by Exponential Smoothing

INTRODUCTION

Forecasting plays a crucial role in the development of plans for the future. It is essential to know for what level of activity one is planning before any commitment of the three process inputs, men, machines, and capital, can be made. Those in charge of planning are faced with many questions. What size must the capital budget be? How large a labor force is needed? What size orders should be placed? What size safety stock should be maintained? What shall be the capacity of the building? The answers to these and many other questions depend upon the forecast of the future level of activity.

MAJOR FORECASTING APPROACHES

These future levels are determined in many ways. One way to categorize the techniques used in forecasting is as follows:

1. Judgmental analysis
2. Econometric forecasting models
3. Time series models

The first technique, a method which has been with us for a long time, relies on the art of human judgment. The next two techniques, on the other hand, are relatively new and rely on the field of statistics. These statistical techniques are by no means a panacea, but they often provide a good basis from which a forecast can be made. In practice, then, statistical analysis is very often combined with judgment in the final stage of the forecasting process.

Consider the problem faced by the production manager of the Splitend Hairdryer Company. Splitend is a company which manufactures several models of consumer and commercial hair dryers. In the past, they have used statistical techniques to forecast demand for their products quite successfully. This month the production manager has learned that the company will promote one of their consumer products by a series of TV commercials. He has in his possession the statistical forecast for next month's demand, and rather than use this figure at face value, he adjusts it upward to what he feels is a reasonable level. In this way he will be assured that the demand for increased units will be met from stock.

In this example, the promotional campaign represents but one of the many imponderables that no statistical forecasting routine can consider. Here as in many other situations some human intervention is necessary.

JUDGMENTAL TECHNIQUES

First we will take a closer look at these judgmental techniques. They might rely heavily upon consumer opinions, customer opinions, surveys to distributors, executive opinions, marketing trials, trade journal surveys, and salesmen's estimates. In the case of consumer opinions, it is not uncommon for companies to send out questionnaires to a randomly selected group of consumers. Intentions to buy and attitudes toward the product can be determined, which in turn can provide valuable inputs to the forecasting process. Customers who have already purchased the product are also useful sources of information. A questionnaire could be attached to the product, and in this way the customer's intention to purchase the product again and his satisfaction with the product can be determined.

Distributors such as retail and wholesale outlets can provide some insight into the pace of current and future sales. Executive opinions are perhaps not the best source of estimates, since this group is far removed from the marketplace. However, careful consideration of their expectations and the reasons behind them can be quite beneficial.

In some cases the risk taken by introducing a new product or changing the present product can be substantial. A marketing trial might be an effective way to sample the effect on overall demand. The trial usually takes place in restricted geographical regions and for a short period of time.

Trade journals are another source of information on future levels of demand. Very often they include surveys and interviews with key people in the industry. Finally, one must consider the closest source of judgmental information: the salesman. In many cases he is the least objective, but it may be possible to correct for forecast bias based on historical evidence—that is, if he is always between 15 and 25 percent high, then correct for this upward bias.

ECONOMETRIC FORECASTING

In econometric forecasting the analyst tries to uncover the cause-and-effect relationship between sales and some other phenomena that are related to sales. For example, an appliance manufacturer might discover that the sales of refrigerators respond to the disposable income of consumers with a 1-month lag. That is, 1 month after a change in disposable income, there is a proportionate change in refrigerator sales. Perhaps there is more than one factor that can explain the level of sales. For example, refrigerator sales might be better explained by looking at disposable income *and* housing starts. The analyst therefore tries to identify those factors that best explain the level of sales for a product.

This process is called *econometric forecasting*, and it utilizes regression and correlation techniques. The objective is to establish a cause-and-effect relationship between changes in the sales level of the product and a set of relevant explanatory variables.

TIME SERIES ANALYSIS

The third method, time series analysis, identifies the historical pattern of demand for the product and projects or *extrapolates* this into the future. For example, we might look at the behavior of refrigerator sales over the past 10 years and from this identify four interacting components that would be used in forecasting the next period's demand. These would include trend, cyclical, seasonal, and irregular components. *Trend* is the long-run historical component of the series. *Seasonal* variations typically have a cycle lasting 1 year and are repetitive. They occur because of such influences as weather and, in the case of automobile purchases, a marketing strategy. *Cycles* are more difficult to explain. They last longer than 1 year and can be attributed to the influence of the business cycle. The *irregular* component represents all those variations which are not caused by the trend, seasonal, and cyclical components.

Three of these demand components can be seen in Fig. 11-1. Here we have quarterly sales data for 4 years. A noticeable straight-line trend is evident over this period. In addition, it can be seen that sales in the second quarter are much higher than this trend, and sales in the fourth quarter are much lower. In fact, the same seasonal pattern persists over the 4-year period. Finally, there is an irregular component that overlies the trend and seasonal patterns.

In this chapter we will focus on one of the more effective techniques of time series analysis known as *exponential smoothing*. First we will develop the simplest exponential smoothing model, one which can be used when the time series exhibits only irregular movements. Then we will develop a model which can be used when trend and irregular components are present. Following this, a model for seasonal and irregular components is generated and, finally,

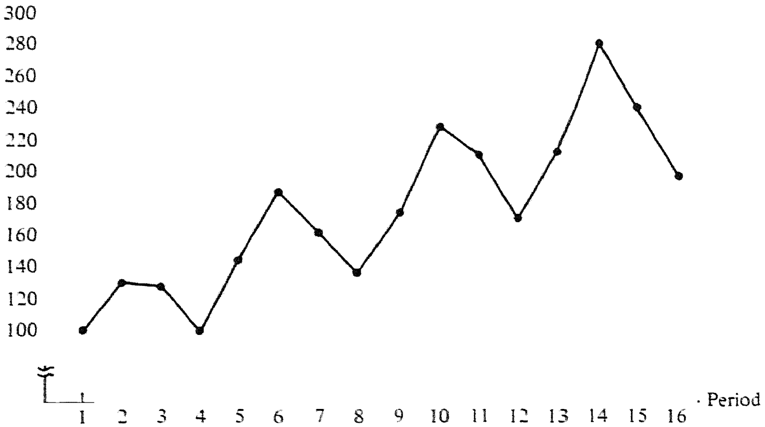


Fig. 11-1 Time series with trend, seasonal, and irregular components.

Fig. 11-2 A series with irregular movements.

a model for trend, seasonal, and irregular components. Having developed models for use under all eventualities, control system concepts are presented whose function it is to monitor and control an operating forecasting system.

A MODEL FOR IRREGULAR MOVEMENTS: SIMPLE EXPONENTIAL SMOOTHING

WEIGHTED AVERAGE

The simplest kind of time series is one in which there is only irregular movement. One such series is shown in Fig. 11-2. If a forecast was necessary, a simple average or even a weighted average could be used. If we did weight the data in such a way that the most recent data had the heaviest weight, our forecasting scheme could respond more effectively to recent changes in the level of the series. Using general notation, a weighted average with six terms can be written as

$$\bar{F} = W_n d_n + W_{n-1} d_{n-1} + W_{n-2} d_{n-2} + W_{n-3} d_{n-3} + W_{n-4} d_{n-4} + W_{n-5} d_{n-5}$$

where $\bar{F}$ = new average to be used as forecast
W_n = weight used in month n
d_n = actual demand in month n

Only one restriction must be made on these weights: their sum must equal 1.

$$\sum_{i=1}^{n} W_i = 1$$

SIMPLE EXPONENTIAL SMOOTHING

This example of a weighted average is very close to exponential smoothing. There are two differences, the first of which is that in an exponentially smoothed average the series continues indefinitely into the past.

$$\bar{F} = W_n d_n + W_{n-1} d_{n-1} + W_{n-2} d_{n-2} + \cdots$$

As in the weighted average, the sum of the weights equals 1.

$$\sum_{i=1}^{n} W_i = 1$$

The second difference is that in exponential smoothing the weights are chosen in a special way. Rather than arbitrarily choosing each W_i, we express the weights as powers of a smoothing constant alpha.

$$\bar{D}_n = \alpha d_n + \alpha(1-\alpha) d_{n-1} + \alpha(1-\alpha)^2 d_{n-2} + \alpha(1-\alpha)^3 d_{n-3} + \alpha(1-\alpha)^4 d_{n-4} + \qquad (11\text{-}1)$$

$\bar{D}_n$ will be used to represent the exponentially smoothed average computed in month n. Given our new weighting scheme, we have only to specify a value of α somewhere between 0 and 1, and all the weights extending into the past are completely specified. In addition, they will decrease with more weight on the most recent months and less and less weight as the months recede into the past. As an example of this, consider a value of .1 for α. The exponentially smoothed function will have the following weights:

$$\bar{D}_n = .1 d_n + .09 d_{n-1} + .081 d_{n-2} + .073 d_{n-3} + .065 d_{n-4} + \cdots$$

More on how to choose this value and the consequence of different values will be given shortly. First, however, it is useful to specify the exponentially smoothed function in a simpler but equivalent form.

$$\bar{D}_n = \bar{D}_{n-1} + \alpha(d_n - \bar{D}_{n-1}) \qquad (11\text{-}2)$$

where $\bar{D}_{n-1}$ = exponentially smoothed average determined last month, used as estimate of this month's demand
d_n = this month's actual demand

We can therefore conclude that the new exponentially smoothed average $\bar{D}_n$ is equal to the old average $\bar{D}_{n-1}$ plus some fraction of the difference between what was predicted and what actually occurred. Said another way, the new average is equal to the old average plus some percent of the error in estimating this month's demand.

Perhaps an illustration can clarify any confusion. Suppose that the forecast made last month for this month's demand was 110. Then $\bar{D}_{n-1} = 110$. The smoothing constant used is $\alpha = .1$, and the actual demand this month was $d_n = 112$. The new smoothed average is therefore:

$$\bar{D}_n = \bar{D}_{n-1} + \alpha(d_n - \bar{D}_{n-1})$$
$$\bar{D}_n = 110 + .1(112 - 110) = 110.2$$

Using this new average, we would forecast next month's demand to be 110.2. Notice that, using this scheme, the data storage problem is quite small. Only last month's smoothed average need be retained. If a conventional weighted average had been used, the demand for each of the past months would have had to be stored.

//// **Proof of shortened form** Before attacking a full-blown example of exponential smoothing, it might be helpful if you were convinced that the shortened form of Eq. 11-2 is equivalent to Eq. 11-1. Repeating Eq. 11-2, we have

$$\bar{D}_n = \bar{D}_{n-1} + \alpha(d_n - \bar{D}_{n-1})$$

By rearranging terms, we can also write the expression as

$$\bar{D}_n = \bar{D}_{n-1} + \alpha d_n - \alpha\bar{D}_{n-1}$$

or

$$\boxed{\bar{D}_n = \alpha d_n + (1 - \alpha)\bar{D}_{n-1}}$$

Since $\bar{D}_{n-1}$ was determined by updating a smoothed estimate *two periods old*, we have

$$\bar{D}_n = \alpha d_n + (1 - \alpha)[\alpha d_{n-1} + (1 - \alpha)\bar{D}_{n-2}]$$

or

$$\bar{D}_n = \alpha d_n + \alpha(1 - \alpha)d_{n-1} + (1 - \alpha)^2\bar{D}_{n-2}$$

Again, since $\bar{D}_{n-2}$ was determined by updating a smoothed estimate *three periods old*, we have

$$\bar{D}_n = \alpha d_n + \alpha(1 - \alpha)d_{n-1} + (1 - \alpha)^2[\alpha d_{n-2} + (1 - \alpha)\bar{D}_{n-3}]$$

or

$$\bar{D}_n = \alpha d_n + \alpha(1 - \alpha)d_{n-1} + \alpha(1 - \alpha)^2 d_{n-2} + (1 - \alpha)^3\bar{D}_{n-3}$$

This substitution process could continue indefinitely, but we can already see the following pattern developing.

$$\bar{D}_n = \alpha d_n + \alpha(1 - \alpha)d_{n-1} + \alpha(1 - \alpha)^2 d_{n-2} + \alpha(1 - \alpha)^3 d_{n-3} + \alpha(1 - \alpha)^4 d_{n-4} + \cdots +$$

It can therefore be concluded that Eqs. 11-1 and 11-2 are equivalent and, furthermore, that the use of Eq. 11-2 simplifies the data storage problem. ////

Example Consider the time series presented in Fig. 11-2. Using Eq. 11-2 and letting $\alpha = .1$ we can compute the smoothed average period by period. Getting started, however, might be difficult, since in the first period we have no previous smoothed average to use in the exponential equation. We might therefore set this smoothed average $\bar{D}_1$ equal to actual demand in that period d_1. Therefore we have $\bar{D}_1 = 8$. In the next period the new updated average is computed from Eq. 11-2 in the following way.

$$\bar{D}_2 = 8 + .1(6 - 8) = 7.8$$

This process continues for period 3, where we have

$$\bar{D}_3 = 7.8 + .1(9 - 7.8) = 7.92$$

These and the remaining smoothed averages are entered in Table 11-1. If we were actually at the end of period 7, we would forecast a demand of 7.666 for period 8.

CHOICE OF THE SMOOTHING CONSTANT α

Now that we have settled upon a forecasting technique, the problem is one of selecting a value for α.

The objective in choosing a value for α is to respond to bona fide changes in level but not to respond to random unexplainable changes, sometimes called *noise*. If α is set at 1.0, the *new* average will be set equal to this period's actual demand. This is equivalent to putting all the weight on this period's level of activity and none on previous periods. Therefore, if there is a large noise component to demand, estimates from one period to the next will fluctuate wildly. On the other hand, if there is a bona fide change in the level of demand, it will be reflected immediately and fully in the new average.

As α is shifted from 1 toward zero, the emphasis shifts from the most recent data more evenly over all of the periods. For low values of α, then, the response to sudden changes becomes sluggish, regardless of whether the

Table 11-1 Simple exponential smoothing example

Period	*Demand d_n*	*Smoothed average $\bar{D}_n$*
1	8	[8]
2	6	7.8
3	9	7.92
4	8	7.928
5	7	7.8352
6	8	7.85168
7	6	7.666512

change is bona fide or just noise. There is no way to determine beforehand which value of α will perform most effectively. Several values must be tried.

Retrospective testing is often employed for the purpose of comparing the forecast ability of different α values. This is done in the following way. First a value of α is chosen. Often that value is close to .1. Using historical data, a forecast for the second period is made by computing a smoothed average from the data available in period 1. Then this forecasted value, together with the actual value in period 2, is plotted on a graph. The process continues until points for all of the periods contained in these data are plotted. Next a different value of α is used to determine new forecasted values. These points are also plotted. Then another value of α might be tried, and so on. Finally it should be possible by examination to determine which value of α tracks the actual demand best.

MODELS FOR TREND AND IRREGULAR MOVEMENTS

There are two models which are used when there is reason to believe that trend and irregular components are present. First we turn to the better-known of the two.

THE SIMPLE EXPONENTIAL SMOOTHING MODEL CORRECTED FOR LAG DUE TO TREND

In this model the first step is to compute the simple exponentially smoothed average, using Eq. 11-2. Consider a firm with the sales record shown in Fig. 11-3. We can see that there is trend, accompanied by a small amount of noise. With $\alpha = .1$ and using Eq. 11-2, the new smoothed average is computed period by period. Again the first period presents a problem. We solve it by letting $\bar{D}_1$ equal actual demand in that period, or $\bar{D}_1 = 12$. In the second period the new smoothed average is computed in the following way:

$$\bar{D}_2 = 12 + .1(17 - 12) = 12.50$$

This figure, the smoothed average for period 2, is *used* as the forecast for demand in period 3 for retrospective testing purposes. Therefore, when we plot the actual and forecast series to determine the effectiveness of the forecasting model, the actual demand, 17, is plotted in period 2 but the smoothed average, 12.50, is plotted in period 3.

The process of computing smoothed averages for periods 2 through 30 continues, and the results are shown in Table 11-2. The actual and computed values are plotted in Fig. 11-3, and upon observation the smoothed series does not seem to be doing its job too well. The forecast series lags actual demand.

Corrections for lag due to trend Now we see the problem that occurs when the simple exponential model is used in the presence of a trend. The exponentially smoothed series will lag behind the actual series. That is, the forecasts

Period	Demand	Period	Demand	Period	Demand	Period	Demand
1	12	9	25	17	54	24	60
2	17	10	33	18	49	25	64
3	19	11	30	19	45	26	70
4	16	12	30	20	60	27	65
5	22	13	36	21	63	28	74
6	22	14	44	22	58	29	73
7	24	15	39	23	55	30	70
8	30	16	44				

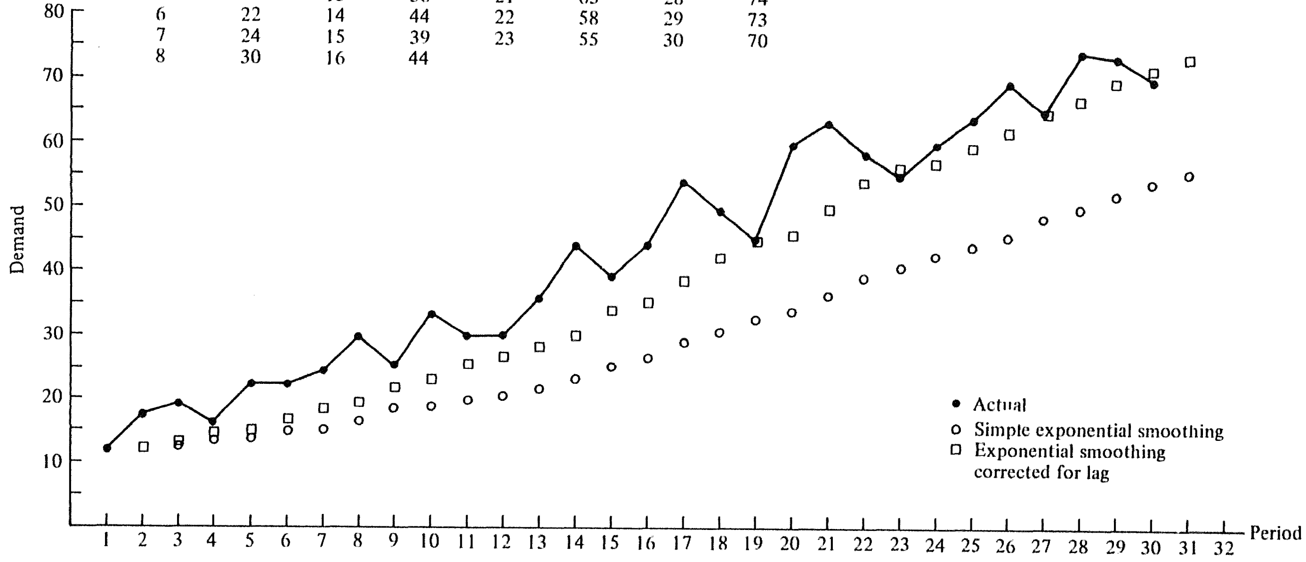


Fig. 11-3 Illustration of simple exponential smoothing and exponential smoothing corrected for lag due to trend.

Table 11-2 Problem which exhibits trend and irregular components

Period	*Demand* d_n	$\bar{D}_n$	*Period*	*Demand* d_n	$\bar{D}_n$
1	12	[12]	16	44	28.26
2	17	12.50	17	54	30.83
3	19	13.15	18	49	32.65
4	16	13.43	19	45	33.88
5	22	14.29	20	60	36.49
6	22	15.06	21	63	39.14
7	24	15.95	22	58	41.03
8	30	17.36	23	55	42.42
9	25	18.12	24	60	44.18
10	33	19.61	25	64	46.16
11	30	20.65	26	70	48.55
12	30	21.58	27	65	50.19
13	36	23.02	28	74	52.57
14	44	25.12	29	73	54.61
15	39	26.51	30	70	56.15

will be consistently low. We therefore must correct for this lag. This is accomplished in the following way.

$$\bar{D}'_n = \bar{D}_n + \{\text{correction for lag due to trend}\}$$

where $\bar{D}'_n$ = new trend-corrected average

It is beyond the scope of this text to derive the trend correction. Therefore, we will write the new trend-corrected average as

$$\bar{D}'_n = \bar{D}_n + \frac{1-\alpha}{}\bar{T}_n \qquad (11\text{-}3)$$

where $\bar{T}_n$ = exponentially smoothed trend

$\bar{T}_n$ is computed using the same approach as in the exponentially smoothed average. First an actual trend is calculated by taking the difference between this period's smoothed demand $\bar{D}_n$ and the previous period's $\bar{D}_{n-1}$.

$$t_n = \bar{D}_n - \bar{D}_{n-1} \qquad (11\text{-}4)$$

In our current example, the actual trend in period 6 would be

$$t_6 = \bar{D}_6 - \bar{D}_5 = 15.06 - 14.29 = .77$$

The next step is to compute the exponentially smoothed trend.

$$\bar{T}_n = \bar{T}_{n-1} + \alpha(t_n - \bar{T}_{n-1}) \qquad (11\text{-}5)$$

The process of smoothing should now be familiar to you. The new smoothed trend is equal to the old smoothed trend $\bar{T}_{n-1}$ plus some percentage α of the difference between the trend that actually occurred and the old estimate. Now we return to the example which we started in Table 11-2.

Example The first step in computing the trend correction is to determine the actual trend for each period. Again the first-period entry presents a problem. There is no value for the exponentially smoothed average in the period preceding the first. One could therefore assume an actual trend of zero for the first period, or make some other reasonable estimate by looking at the graph in Fig. 11-3. Choosing the first approach, the actual trend in the first period is estimated as $t_1 = 0$. The remaining values for t_n have been calculated and can be found in Table 11-3.

The first value for $\bar{T}_n$ is also a problem. Having set the actual trend in that first period of zero, we also set the smoothed trend $\bar{T}_1 = 0$. The remaining

Table 11-3 Exponential smoothing corrected for lag due to trend

Period	d_n	$\bar{D}_n$	t_n	$\bar{T}_n$	$\bar{D}'_n$	$\bar{D}'_{n,1}$
1	12	12.00	0.0	0.0	12.00	12.00
2	17	12.50	.50	.05	12.95	13.00
3	19	13.15	.65	.11	14.14	14.25
4	16	13.43	.28	.12	14.58	14.70
5	22	14.29	.85	.20	16.09	16.29
6	22	15.06	.77	.25	17.37	17.62
7	24	15.95	.89	.32	18.84	19.16
8	30	17.36	1.40	.42	21.22	21.64
9	25	18.12	.76	.46	22.29	22.75
10	33	19.61	1.48	.56	24.70	25.26
11	30	20.65	1.03	.61	26.16	26.77
12	30	21.58	.93	.64	27.38	28.02
13	36	23.02	1.44	.72	29.54	30.26
14	44	25.12	2.09	.86	32.88	33.74
15	39	26.51	1.38	.91	34.74	35.65
16	44	28.26	1.74	.99	37.24	38.23
17	54	30.83	2.57	1.15	41.23	42.38
18	49	32.65	1.81	1.22	43.64	44.86
19	45	33.88	1.23	1.22	44.89	46.11
20	60	36.49	2.61	1.36	48.75	50.11
21	63	39.14	2.65	1.49	52.56	54.05
22	58	41.03	1.88	1.53	54.80	56.33
23	55	42.42	1.39	1.51	56.08	57.59
24	60	44.18	1.75	1.54	58.05	59.59
25	64	46.16	1.98	1.58	60.43	62.01
26	70	48.55	2.38	1.66	63.53	65.19
27	65	50.19	1.64	1.66	65.16	66.82
28	74	52.57	2.38	1.73	68.18	69.91
29	73	54.61	2.04	1.76	70.50	72.26
30	70	56.15	1.53	1.74	71.83	73.57

values for $\bar{T}_n$ are computed, using Eq. 11-5. For example, the smoothed trend in period 2 is

$$\bar{T}_2 = \bar{T}_1 + \alpha(t_2 - \bar{T}_1)$$
$$\bar{T}_2 = 0 + .1(.50 - 0)$$
$$\bar{T}_2 = .05$$

The process continues and the results can be seen in Table 11-3.

Next, the new average $\bar{D}'_n$ is computed using Eq. 11-3. In period 1 the calculation is

$$\bar{D}'_1 = \bar{D}_1 + \frac{1-\alpha}{\alpha}\bar{T}_1$$
$$\bar{D}'_1 = 21$$

For period 2 we have

$$\bar{D}'_2 = 12.5 + 9(.05) = 12.95$$

Again the process continues for periods 3 through 30, and the results are shown in Table 11-3.

The value $\bar{D}'_n$ represents a smoothed average, computed in period n, which has been corrected for lag due to trend. Since the presence of a trend has been acknowledged, it seems reasonable to use this trend in making a forecast. Therefore the forecast k periods hence, which we will denote by $\bar{D}'_{n,k}$, can be computed in the following way.

$$\bar{D}'_{n,k} = \bar{D}'_n + k\bar{T}_n$$

For example, the forecast for period 2, made in period 1, is

$$\bar{D}'_{1,1} = \bar{D}^1_1 + \bar{T}_1$$
$$\bar{D}'_{1,1} = 12.0 + 0.0$$
$$\bar{D}'_{1,1} = 12$$

The forecast for period 3, made in period 2, is

$$\bar{D}'_{2,1} = 12.95 + .05$$
$$\bar{D}'_{2,1} = 13.00$$

The computational process continues, and the results are shown in Table 11-3.

The forecast, which is graphed in Fig. 11-3, starts off by lagging the actual series, but it adjusts slowly and eventually proves to track the actual series quite well. The period of adjustment is necessary since we made some assumptions about those first values of $\bar{D}_1$, t_1, and $\bar{T}_1$. Therefore, in any exponentially smoothed model, it is wise to provide for this period of adjustment before using the results for planning purposes. If historical data are available, this adjustment process can be executed on the computer or paper very much as we did it in this example. But in the case of a new product with no historical data, the trend-corrected average $\bar{D}'_n$ cannot be used for several periods into the future until such time as the exponentially smoothed function can be shown to be tracking well. If a larger value of α was used initially, the adjustment process would have taken less time, but the response to random variations about the bona fide trend would also have been greater, probably resulting in greater forecast errors. The only way to be sure of selecting the best value of α is to make several trials.

//// ANOTHER MODEL

There is another more versatile model which can be used in the presence of a trend and irregular components. In fact, it will serve as the basis of our more complex models in the following sections. Using the same general notation, we have

$$\bar{D}_n = (\bar{D}_{n-1} + \bar{T}_{n-1}) + A[d_n - (\bar{D}_{n-1} + \bar{T}_{n-1})]$$
$$\bar{T}_n = \bar{T}_{n-1} + B[(\bar{D}_n - \bar{D}_{n-1}) - \bar{T}_{n-1}]$$

and

$$\bar{D}_{n,k} = \bar{D}_n + k\bar{T}_n$$

Example A very simple example is shown in Table 11-4. It illustrates how this model would perform in predicting a linear series which is increasing by a constant amount. Letting $A = .1$ and $B = .1$, and setting the initial values of $\bar{D}_1$ and $\bar{T}_1$ at 10 and 2, respectively, we can compute the smoothed average, smoothed trend, and forecast for period 2 in the following way.

$$\bar{D}_2 = (10 + 2) + .1(12 - 12) = 12$$
$$\bar{T}_2 = 2 + .1[(12 - 10) - 2] = 2$$
$$\bar{D}_{2,1} = 12 + 2 = 14$$

In a similar fashion the forecast for the next three periods can be computed. Given such well-behaved data, it is clear that this model does a perfect job.

If this model had been used for the data presented in Table 11-2, it would have tracked the series equally as well as the simple exponential model corrected for lag, once both of them had passed their adjustment period. ////

Table 11-4 Example with trend and irregular components using second model

Period i	*Demand d_i*	$\bar{D}_n$	$\bar{T}_n$	$\bar{D}_{n,1}$
1	10	[10]	2	12
2	12	12	2	14
3	14	14	2	16
4	16	16	2	18
5	18	18	2	20

//// A BETTER WAY OF COMPARING SMOOTHING CONSTANTS

There is a more explicit way of comparing one smoothing constant with another. The variance of the error can be estimated by

$$S_e^2 = \frac{1}{N-1} \sum^{N} (d_n - \bar{D}'_{n-1,1})^2$$

The forecast error in each period is defined as the amount by which the exponentially smoothed forecast $\bar{D}^1_{n-1,1}$ differs from the actual demand in that period d_n. We then square this difference and sum the differences for all the observations. Dividing by $N - 1$, we have the variance of the error terms. The best forecast will have the least amount of variance S_e^2. Notice that in the process of getting a measure for the variability of error terms, positive and negative deviations are weighted equally and extreme deviations are weighted more heavily than small deviations. This means that large forecast errors are considered much more costly than small ones. In most cases this is intuitively acceptable. ////

A MODEL FOR SEASONAL AND IRREGULAR MOVEMENTS

If there is evidence that seasonal factors influence the level of demand, seasonal adjustments should be made to the forecast. These seasonal effects can be either multiplicative or additive. If the amplitude of this component is independent of the level of sales, an additive model should be used. Therefore, if January's demand is determined as 10 units above the average level of demand, 10 units are added to the January forecast each year. In a multiplicative model, the amplitude is proportional to the level of demand. For example, January's demand might be 12 percent above the average. The multiplicative model is more appropriate for most applications, and it will therefore be the basis for our formulation. The first model developed will take into account seasonal and irregular components, and a later model will consider seasonal and irregular as well as trend components.

When seasonal components are present, the data must first be deseasonalized. For example, if January's actual sales rate is 110 and the seasonal factor is 1.10 (January's sales are 10 percent above average), the deseasonalized figure for the month is 110/1.10 = 100. In general terms, we divided the actual sales d_n for month n by a seasonal factor which was computed for the corresponding month in the previous year. We will call the seasonal factor

$$\bar{F}_{n-L}$$

where L = number of periods in seasonal cycle (12 in a yearly cycle and 4 in a quarterly cycle, etc.)

and the subscript therefore represents the period, a season ago, in which the seasonal factor was computed.

We can now write an exponentially smoothed deseasonalized average as

$$\bar{D}_n = \bar{D}_{n-1} + A\left(\frac{d_n}{F_{n-L}} - \bar{D}_{n-1}\right)$$

where $\bar{D}_{n-1}$ = deseasonalized average computed last period
$\dfrac{d_n}{\bar{F}_{n-L}}$ = this period's deseasonalized actual demand
A = smoothing constant

The seasonal factor $\bar{F}_{n-L}$ is based on the pattern of seasonal variation that has occurred L periods ago, $2L$ periods ago, $3L$ periods ago, and so on. For example, the seasonal factor for this January would be based on some weighted average of the seasonal factor last January, the January preceding it, and the one preceding that, and so on. We have already discovered an efficient way of computing and updating a weighted average extending into the past—exponential smoothing.

An exponentially smoothed function for updating seasonal factors can be written as

$$\bar{F}_n = \bar{F}_{n-L} + B\left(\frac{d_n}{\bar{D}_n} - \bar{F}_{n-L}\right)$$

where $\bar{F}_n$ = new exponentially smoothed seasonal factor (which will be stored and used to deseasonalize demand in the corresponding season next year)
$\bar{F}_{n-L}$ = exponentially smoothed seasonal factor computed during corresponding season last year

The fraction $\dfrac{d_n}{\bar{D}_n}$ is the level of actual demand divided by the new updated average computed this period. This represents the degree by which this period's actual demand exceeds the deseasonalized new average. It is, then, the actual seasonal factor for this period. Therefore, the exponential function can be interpreted as follows: the new updated seasonal factor is equal to the old seasonal estimate $\bar{F}_{n-L}$ plus some fraction B of the difference between the actual seasonal factor this period and the old estimate. The smoothing factor B is usually different from the smoothing factor A. More on the choice of these values will be given later.

Since the smoothed average demand is deseasonalized, it is necessary to seasonalize it when making a forecast. Therefore, if we were making a forecast for the next period $n + 1$, we would have

$$\bar{D}_{n,1} = \bar{D}_n\bar{F}_{n-L+1}$$

where $\bar{D}_{n,1}$ = forecast for next period
$\bar{D}_n$ = deseasonalized smoothed demand for period n
$\bar{F}_{n-L+1}$ = seasonal factor applying to next period's forecast and computed $n - L + 1$ periods ago

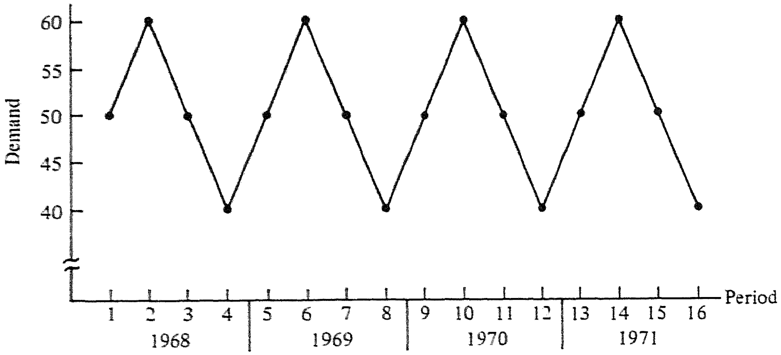


Fig. 11-4 Seasonal example.

Example Consider the time series of quarterly data shown in Fig. 11-4. To illustrate the use of this seasonal model, we will once again start with the data given in period 1 and compare the forecast with the actual to see how well the seasonal model tracks the historical series.

The first problem is that of initial values. As before, we could set $\bar{D}_1$ equal to d_n. The seasonal factors for the four preceding periods could be set at 1. Instead we choose to assume initial values which follow an obvious seasonal pattern. These values are given at the top of Table 11-5. The smoothing

Table 11-5 Seasonal forecast

n	d_n	$\bar{D}_n$	$\bar{F}_n$	$\bar{D}_{n,1}$
			1.0	
			1.2	
			1.0	
		50	.8	
1	50	50	1.0	60
2	60	50	1.2	50
3	50	50	1.0	40
4	40	50	.8	50
5	50	50	1.0	60
6	60	50	1.2	50
7	50	50	1.0	40
8	40	50	.8	50
9	50	50	1.0	60
10	60	50	1.2	50
11	50	50	1.0	40
12	40	50	.8	50
13	50	50	1.0	60
14	60	50	1.2	50
15	50	50	1.0	40
16	40	50	.8	50

constants are set at $A = .1$ and $B = .2$. Therefore the first deseasonalized average is

$$\bar{D}_1 = 50 + .1\left(\frac{50}{1.0} - 50\right) = 50$$

and the updated seasonal factor is

$$\bar{F}_1 = 1.0 + .2(\tfrac{50}{50} - 1) = 1.0$$

The forecast for next period is

$$\bar{D}_{1,2} = 50(1.2) = 60$$

The second deseasonalized factor is

$$\bar{D}_2 = 50 + .1\left(\frac{60}{1.2} - 50\right) = 50$$

and the updated seasonal factor is

$$\bar{F}_2 = 1.2 + .2(\tfrac{60}{50} - 1.2) = 1.2$$

The forecast for the third period is

$$\bar{D}_{2,1} = 50(1.0) = 50$$

These calculations are carried out for the remaining 14 periods in the time series. The results can be seen in Table 11-5. It is clear that the seasonal model does a perfect tracking job in this rather absurdly simple problem. Next we will introduce a linear trend and irregular components as well as seasonal variations.

A MODEL FOR TREND, SEASONAL, AND IRREGULAR MOVEMENTS

If in addition to irregular and seasonal components a trend is present, the expression for deseasonalized demand becomes

$$\bar{D}_n = (\bar{D}_{n-1} + \bar{T}_{n-1}) + A\left[\frac{d_n}{\bar{F}_{n-L}} - (\bar{D}_{n-1} + \bar{T}_{n-1})\right]$$

where $\bar{T}_{n-1}$ = smoothed trend factor computed last period

The expression in the first parenthesis then represents last period's deseasonalized average plus a trend correction. It is therefore a deseasonalized and trend-corrected average for the present period. To this is added some fraction of the amount by which the deseasonalized actual demand exceeds the trend-corrected and deseasonalized old average.

How do we get the trend factor $\bar{T}_n$? By exponentially smoothing! Now you are catching on.

$$\bar{T}_n = \bar{T}_{n-1} + C[(\bar{D}_n - \bar{D}_{n-1}) - \bar{T}_{n-1}]$$

The new smoothed trend factor is equal to the old trend factor plus some fraction C of the difference between the actual trend in the present period and the old trend. The actual trend is computed as the difference between this period's smoothed demand and the old smoothed demand $\bar{D}_n - \bar{D}_{n-1}$. The smoothing constant C does not have to be the same as the ones for the demand and seasonal constants.

The seasonal factor is computed just as before.

$$\bar{F}_n = \bar{F}_{n-L} + B\left(\frac{d_n}{\bar{D}_n} - \bar{F}_{n-L}\right)$$

Finally, to forecast for the next period, using this latest model, we have

$$\bar{D}_{n,1} = (\bar{D}_n + \bar{T}_n)\bar{F}_{n-L+1}$$

where $\bar{D}_n + T_n$ = deseasonalized but trend-corrected estimate for next period and $\bar{F}_{n-L+1}$ is the seasonal factor which is multiplied by the trend-corrected estimate to give us a seasonalized demand estimate. Let us turn to an example where these ideas can be illustrated.

Example Consider the quarterly data presented in Fig. 11-1. These data are repeated in Table 11-6. Initial values for $\bar{D}_n$, $\bar{F}_n$, and $\bar{T}_n$ are determined by estimation. The values of A, B, and C are set at .20, .10, and .30, respectively. The sequence of calculations is to first compute the deseasonalized smoothed

Table 11-6 An example with trend, seasonal, and irregular components

Period	d_n	$\bar{D}_n$	$\bar{F}_n$	$\bar{T}_n$	$\bar{D}_{n,1}$
			1.00		
			1.20		
			1.00		
		90	.80	10	
1	100	100	1.00	10	132.00
2	128	109.333	1.1971	9.8	119.133
3	124	120.107	1.0032	10.092	104.159
4	100	129.159	.7974	9.7801	138.939
5	142	139.551	1.0017	9.9637	178.980
6	185	150.521	1.2003	10.2655	161.307
7	160	160.526	1.0026	10.1873	136.130
8	134	170.178	.7964	10.0270	180.522
9	173	178.704	.9984	9.5765	225.988
10	228	188.616	1.2011	9.6771	198.806
11	211	200.725	1.0075	10.4068	168.150
12	169	211.345	.7967	10.4708	221.458
13	216	220.723	.9964	10.1428	277.299
14	280	231.315	1.2021	10.2777	243.393
15	240	240.919	1.0063	10.0757	199.979
16	195	249.745	.7951	9.7007	258.514

demand $\bar{D}_n$, then the smoothed seasonal factor $\bar{F}_n$, then the smoothed trend factor $\bar{T}_n$, and finally the forecast for next period $\bar{D}_{n,1}$. For period 1 we have

$$\bar{D}_1 = (90 + 10) + .2\left[\frac{100}{1.00} - (90 + 10)\right]$$

$$\bar{D}_1 = 100$$

$$\bar{F}_1 = 1.00 + .1[\tfrac{100}{100} - 1.00]$$

$$\bar{F}_1 = 1.00$$

$$\bar{T}_1 = 10 + .3[(100 - 90) - 10]$$

$$\bar{T}_1 = 10$$

$$\bar{D}_{1,1} = (100 + 10)1.20$$

$$\bar{D}_{1,1} = 132.00$$

For period 2 we have

$$\bar{D}_2 = (100 + 10) + .2\left[\frac{128}{1.2} - (100 + 10)\right]$$

$$\bar{D}_2 = 109.333$$

$$\bar{F}_2 = 1.2 + .1\left[\frac{128}{109.333} - 1.2\right]$$

$$\bar{F}_2 = 1.1971$$

$$\bar{T}_2 = 10 + .3[(109.33 - 100) - 10]$$

$$\bar{T}_2 = 9.8$$

$$\bar{D}_{n,1} = (109.333 + 9.8)1.00$$

$$\bar{D}_{n,1} = 119.133$$

The calculations for the remaining periods can be found in Table 11-6. The figures are then plotted in Fig. 11-5, and it can be seen that the forecast is very close to the actual.

INITIAL VALUES OF *A*, *B*, AND *C*

In many cases the objective of the forecasting model is to minimize the forecast errors. Therefore in most situations that model which minimizes the standard deviation of forecast errors is chosen as the best. This criterion has little more to justify it than mathematical expediency, in spite of its widespread use. The proper criterion should be related to the consequence of forecast errors upon the overall objective.[1] To deal with the problem at this aggregate level is quite difficult and accounts for the reliance on the minimization of the standard deviation of forecast errors.

---

[1] Arthur M. Goeffrion, "A Summary of Exponential Smoothing," *Journal of Industrial Engineering*, vol. 13, no. 4, July–August 1967, pp. 223–226.

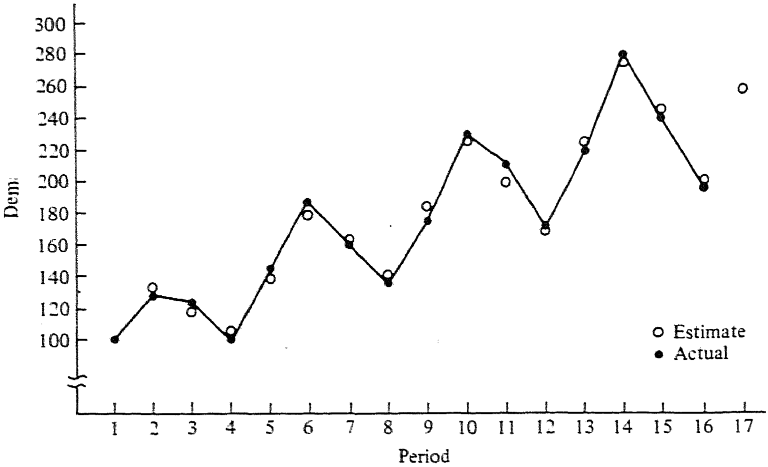


Fig. 11-5 Trend, seasonal, and irregular problem: actual and forecast.

One set of exponential smoothing constants A, B, and C is likely to have a different standard deviation of forecast errors from that of another set. Therefore many combinations should be tried before choosing one set. Winters[2] used this technique in determining A, B, and C for an exponential smoothing model. He tried all possible combinations of the weights 0, 0.2, 0.4, 0.6, 0.8, and 1.0 for each of the three smoothing factors. For each of the 216 sets he calculated a standard deviation of forecast errors. Near the set which was associated with the minimum standard deviation he used a finer grid of values. He concluded that the standard deviation was generally insensitive to changes in the smoothing constants in the vicinity of their minimum.

GENERALITY OF MODELS

The models in this chapter have been developed according to the following progression:

1. Simple exponential smoothing
2. Simple exponential smoothing corrected for trend
3. Exponential smoothing with seasonal components
4. Exponential smoothing with trend and seasonal components

There is no intended implication by this progression that model 2 is more general than model 1, that model 3 is more general than model 2, etc. It therefore does not hold that if the model builder is uncertain as to the presence of trend and seasonal, he is better off to use model 4. If in fact he follows this impulse and no trend or seasonal components are *actually* present, the model will be reacting to certain sequences of noise and interpreting them as influences on the seasonal and trend factors. The correction made by these seasonal and trend factors will be inappropriate and the errors so generated will be greater than if the simple exponential smoothing model were used.

---

[2] P. R. Winters, "Forecasting Sales by Exponentially Weighted Moving Averages," O.N.R. Research Memorandum no. 62, Carnegie Institute of Technology, March 1959.

The problem, then, is not simply a matter of finding the right smoothing constants but one of also choosing the right model.

A CONTROL SYSTEM

We can develop a system to monitor the magnitude of the difference between the actual and forecast values. By observing the magnitude of this error we can determine when the system is not doing its job, or when the system is "out of control."

MONITORING THE ERROR SIGNAL

First, if we have selected the model correctly, the forecast errors should average out to approximately zero. There are two ways in which we could check this. A cumulative sum of the forecast errors could be monitored; when the sum gets too great, the system would be considered out of control and the appropriate form of intervention would take place. Brown,[3] however, suggests that we develop a smoothed error estimate $\bar{Z}_n$.

$$\bar{Z}_n = \bar{Z}_{n-1} + E(e_n - \bar{Z}_{n-1})$$

where e_n = actual error in period n and can be calculated by subtracting the prediction for this period, made one period ago, from the actual this period:

$$e_n = d_n - \bar{D}_{n-1,1}$$

E = smoothing constant, probably different from A, B, and C

The smoothed error can therefore be interpreted as the old smoothed error $\bar{Z}_{n-1}$ plus some fraction E of the amount by which the actual error exceeds the old smoothed error.

Initial values must be determined for the actual error e_1 and the smoothed error $\bar{Z}_1$. Since the average error is expected to be zero, setting both at zero in the first period seems reasonable. If the smoothing constant for our example is set at $E = .2$, the actual error in period 2 is

$$e_2 = 128 - 132 = -4$$

and the smoothed error is

$$\bar{Z}_2 = 0 + .2(-4 - 0) = -.8$$

The remaining calculations are shown in Table 11-7.

---

[3] Robert G. Brown, *Decision Rules for Inventory Management,* Holt, Rinehart & Winston, Inc., New York, 1967, p. 152.

Table 11-7 A control system for monitoring the error signal

n	e_n	$\bar{Z}_n$	$\overline{\text{MAD}}_n$	$\dfrac{\lvert\bar{Z}_n\rvert}{\sigma_{\bar{Z}n}}$
1	0	0	3.9020	0
2	−4.0	− .8	3.9216	.8294
3	4.8667	.3333	4.1106	.3297
4	−4.1589	− .5651	4.1203	.5576
5	3.0610	.1601	3.9084	.1665
6	6.0197	1.3320	4.3307	1.2505
7	−1.3073	.8041	3.7260	.8774
8	−2.1305	.2172	3.4069	.2592
9	−7.5216	−1.3305	4.2298	1.2788
10	2.0125	− .6619	3.7863	.7107
11	12.1938	1.9092	5.4678	1.4196
12	.8498	1.6973	4.5442	1.5185
13	−5.4584	.2662	4.7271	.2289
14	2.7015	.7532	4.3220	.7085
15	−3.3929	− .0759	4.1361	.0747
16	−4.9789	−1.0565	4.3047	.9979

HELPING THE LEARNING PROCESS

Demand changes, and it is precisely for this reason that we smooth demand each period to adapt the new estimates to changes in demand. The smoothing constants A, B, and C control the rate at which the model learns. However, if rapid changes take place, it might be necessary to "help" this learning process along. Take the case where a new TV set has been introduced into a product line. This set uses a popular transistor which is used in many other products which the firm produces. Forecasts for the unit have, in the past, been made from an exponential smoothing model. Upon introduction of the new set, the underlying demand pattern for this transistor is expected to change sharply. It seems inefficient to allow the normal reactions of the exponential mode to adapt to this increase. It would be much more effective if the values of A, B, and C were increased temporarily until the new demand pattern was being tracked satisfactorily.

In most cases it is *not* predictable just when and why underlying demand pattern may change. What we therefore need is to monitor the amount by which the actual differs from the forecast demand—the forecast error. When a significant error is detected, the system should allow for a temporary change of A, B, and C. Next we turn to identifying just when that moment for intervention has arrived.

SETTING CONTROL LIMITS

In a preceding section the concept of the smoothed error signal Z_n was developed. Now we turn to the estimation of its *variability*, and we start by estimating the mean absolute deviation MAD. Again the procedure is to compute a smoothed average.

$$\overline{\text{MAD}}_n = \overline{\text{MAD}}_{n-1} + E(|e_n| - \overline{\text{MAD}}_{n-1})$$

where $\overline{\text{MAD}}_{n-1}$ = smoothed mean absolute deviation computed last period
E = same smoothing constant used in computing smoothed error
$|e_n|$ = absolute value of actual forecast error in period n

The computations for $\overline{\text{MAD}}_n$ are shown in Table 11-7. First, the initial value for $\overline{\text{MAD}}_1$ must be determined. A reasonable approach would be to take the average of the error terms e_n for the first 6 periods for which data were available. Computing this value for periods 2 through 7, we estimate $\overline{\text{MAD}}_1$ to be 3.902. Computing the $\overline{\text{MAD}}_n$ value for the second period, we have

$$\overline{\text{MAD}}_2 = 3.9020 + .2(4.0 - 3.9020)$$

$$\overline{\text{MAD}}_2 = 3.9216$$

If the forecast is accurate, the average forecast error should be zero, and the smoothed error $\bar{Z}_n$ should fluctuate around zero. Brown[4] states, however, that even with excellent forecasts, the *smoothed* error will not be zero. He shows that the standard deviation of the smoothed error $\sigma_{\bar{Z},n}$ in period n appears to be

$$\sigma_{\bar{Z},n} = .55\sqrt{E}\,(\overline{\text{MAD}}_n)$$

We can therefore make statements such that for a suitable value K, if the ratio $|\bar{Z}_n|/\sigma_{\bar{Z}_n}$ is less than K, the smoothed error is small enough to be considered a random fluctuation. If E is small, the distribution of the ratio $|\bar{Z}_n|/\sigma_{\bar{Z},n}$ is nearly normal, and the normal table can be used to set the control limit K. For example, we can say that 95 percent of the time observations from a process with mean $\bar{D}_n$ will be such that $|\bar{Z}_n|/\sigma_{\bar{Z},n} \leq 2$. Said another way, 95 percent of the time the smoothed error will be within 2 standard deviations of the mean. Therefore, if we set the control limit at 2 standard deviations, only 5 percent of the time would we get a larger variation, when, in fact, the system is actually in control. This is, of course, the risk we take in any statistically oriented control system: A certain percent of the time, an erroneous decision is made. Action is taken to rectify the situation when, in fact, no action is really necessary. Most of the time, however, when an error signal exceeds this limit, there is a bona fide problem and human intervention is necessary.

The form that the human intervention takes is to readjust A, B, and C upward and thereafter closely monitor the ability of the system to track the actual demand until the system is once again under control. Another form that intervention might take is to reexamine the time series in an effort to determine whether another model might be more appropriate. For example,

---

[4] Ibid., p. 153.

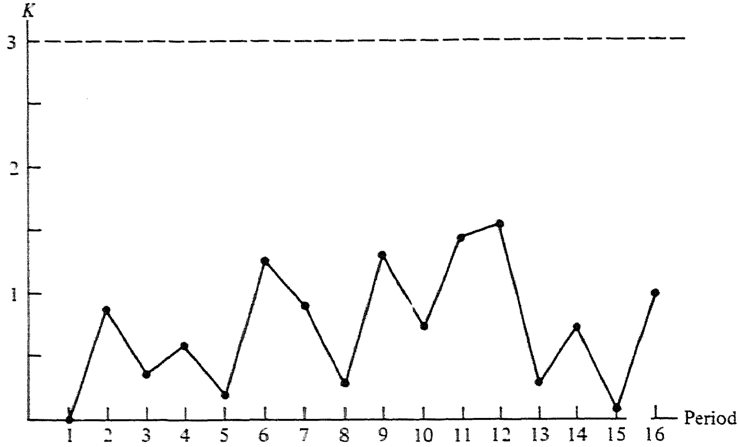


Fig. 11-6 Control chart for forecast error.

if a simple exponential model is being used, an out-of-control signal may be caused by the appearance of a trend. The introduction of the exponential smoothing model corrected for trend would almost certainly solve the problem.

Getting back to the example, a 3σ or 99 percent limit will be set. Therefore, any value of the ratio in excess of $K = 3$ will precipitate a readjustment of A, B, and C for a number of periods until the new demand pattern is tracked satisfactorily. In the first period the ratio is:

$$\frac{|\bar{Z}_1|}{\sigma_{\bar{Z},1}} = \frac{0}{.55\sqrt{.2}\,(3.9020)} = 0$$

and in the second period we have

$$\frac{0.8}{0.55\sqrt{.2}\,(3.9216)} = 0.8294$$

The calculations are continued in Table 11-7. It can be seen that the computed ratio never exceeds the critical value of 3. Therefore, the system is in control. A more graphic way to portray the control process is by means of a control chart. Figure 11-6 represents such a chart with a dashed line drawn at $K = 3$ representing the control limit.

THE CHOICE OF K

A reasonable approach to setting up a control system is to segment the items predicted. For example, category I might include the fastest-moving 5 percent of the items that account for half of the dollar volume of demand. At the other extreme you might have category III, which accounts for half of the items in inventory but represents a small percent of dollar volume. All other items would be classified as category II. We would therefore set our K limits accordingly. For category I items we would like to know as soon as possible if the system is out of control, and therefore would be willing to pay the price for false reports or situations when there is no real change in demand but where a random variation exceeds the K level. Therefore we might set category I at 2.5 to 3.5 standard deviations. Category III might be set at 3.5 to 4.5 and category II at 3.0 to 4.0. If too many out-of-control signals are generated to be dealt with effectively, the threshold levels K for categories I, II, and III can be increased.

//// UNCERTAINTY OF THE FORECAST

The forecast that is made for some future period, say next period $\bar{D}_{n,1}$, is a point estimate. For example, we estimate that the demand next period will be 115. However, the very fact that forecast errors have been observed shrouds the point estimate with uncertainty. The uncertainty of the forecast can be communicated by establishing a confidence interval around the point estimate or mean expected demand. We already have a measure of the variability of forecast errors in the $\overline{\text{MAD}}$. The relationship between the standard deviation of forecast errors and $\overline{\text{MAD}}$ is

$$S_{e,n} = 1.25\ \overline{\text{MAD}}_n$$

If we make the fairly reasonable assumption that forecast errors are normally distributed, the normal table can be used to determine confidence intervals. For example, if $\overline{\text{MAD}}_n = 4$,

$$S_{e,n} = 1.25(4) = 5.0$$

and the statement could be made that there is a 95 percent chance that the interval from $115 \pm 2(5)$ or 105 to 125 would include the true demand in the next period. ////

SUMMARY

Exponential smoothing is not the most sophisticated forecasting technique developed. However, it is frequently very effective. This is especially true for the thousands of items in inventory that must be forecast on a routine basis. If the manpower were available to do the forecasting job manually, the cost would be very high and the time devoted to preparing a forecast for each item would likely be small. By comparison, then, a computer system based on exponential smoothing can perform quite effectively. Although mechanized, the system would not go without human intervention. For when the smoothed forecast error becomes too large, an out-of-control signal would be generated, and this, in turn, would be followed by readjustments to the smoothing constants and, if necessary, by operator scrutinization. On some occasions human intervention may be warranted even in the absence of out-of-control signals. For example, the introduction of new products, acts of competitors, and acts of nature all may precede changes in the underlying time series. This change can be anticipated by a temporary increase in the smoothing constants until the model tracks the new pattern effectively.

Setting up an exponential smoothing system is a task of considerable magnitude. First, the type of model must be selected—simple exponential smoothing, simple exponential smoothing corrected for trend, exponential smoothing with seasonal components, or smoothing with seasonal and trend. Second, the initial values for the smoothing constants must be determined, then quite a bit of retrospective testing is involved where values of the smoothing constants are varied and the consequence of these values over the historical series is compared. Finally control limits must be set. Needless to say, a computer is an absolute necessity.

QUESTIONS

Q11-1. What is a weighted average, and why is it unacceptable when a trend is present?

Q11-2. Compute the first six coefficients for the simple exponential smoothing formula (Eq. 11-1) when $\alpha = .30$. Is this a weighted average? How far into the distant past do these coefficients extend, and what does this mean?

Q11-3. What are the consequences of α values of 0 and 1?

Q11-4. How do you know that choice of α is the best possible one?

Q11-5. What is the consequence of using the simple exponential formula in the presence of a seasonal cycle?

Q11-6. What is the consequence of using the simple exponential formula in the presence of a trend?

Q11-7. What steps are taken to adjust the simple exponential forecast for lag due to trend?

Q11-8. If in doubt, always use the most general model, exponential smoothing corrected for trend and seasonal components. True or false? Explain why.

Q11-9. What mechanism can be employed to monitor the forecasting system? If it is determined that the system is "out of control," what should be done?

Q11-10. Explain the process of establishing a control chart for the smoothed error signal. How are the control limits set? What is the consequence of setting these limits too wide? Of setting them too narrow?

Q11-11. How should the smoothing constants A, B, and C be chosen?

Q11-12. How should the initial seasonal factors $\bar{F}_{n-L}$ be chosen? What would be the consequence if each one of them was set at 1 when in fact a seasonal cycle was evident?

Q11-13. Derive an expression for the deseasonalized smoothed average for the case of a seasonal series with an additive seasonal component.

Q11-14. Why might the value of K in a monitoring system vary among different inventory items?

Q11-15. Derive a smoothing model for a trend and an additive seasonal component.

PROBLEMS

P11-1. Using simple exponential smoothing techniques, determine a forecast for periods 2 through 20. Graphically compare your forecast with the actual. Let $\alpha = .2$ and $\bar{D}_1 = 130$.

Period	*Actual*	*Period*	*Actual*
1	130	11	134
2	140	12	144
3	148	13	133
4	138	14	159
5	130	15	114
6	141	16	140
7	136	17	148
8	160	18	157
9	145	19	151
10	130	20	140

//// **P11-2.** Repeat Prob. 11-1 for an $\alpha = .5$. Compute the variance of the error for each alternative. Which one performs best? ////

P11-3. Write a computer program for the simple exponential smoothing formula and use it to solve Prob. 11-1.

P11-4. Using simple exponential smoothing techniques, determine a forecast for periods 2 through 20, for which actual figures appear below. Graphically compare your forecast with the actual. Let $\alpha = .1$ and $\bar{D}_1 = 211$.

Period	*Actual*	*Period*	*Actual*
1	211	11	181
2	196	12	184
3	190	13	211
4	180	14	181
5	220	15	230
6	219	16	200
7	200	17	190
8	209	18	180
9	190	19	180
10	195	20	197

P11-5. Using the simple exponential smoothing formula with a correction for lag, compute a forecast for the following series. Let $\alpha = .2$, $t_1 = 10$, $\bar{T}_1 = 10$, and $\bar{D}_1 = 20$.

Period	*Actual*	*Period*	*Actual*
1	60	5	100
2	70	6	110
3	80	7	120
4	90	8	130

What conclusions can be drawn about this model?

P11-6. Using the simple exponential smoothing formula with a correction for lag due to trend, compute a forecast of the following series for periods 2 through 21. Let $\alpha = .2$, $\bar{D}_1 = 13$, $t_1 = 3$, and $\bar{T}_1 = 3$.

Period	*Actual*	*Period*	*Actual*
1	25	11	45
2	32	12	53
3	41	13	72
4	27	14	58
5	38	15	66
6	34	16	68
7	37	17	71
8	45	18	79
9	70	19	77
10	42	20	87

Graph your results and compare the forecast with the actual.

P11-7. For the following time series, find the values of $\bar{D}_1$, t_1, and $\bar{T}_1$ such that the simple exponential smoothing model corrected for lag will perfectly track the data from the beginning. Let $\alpha = .2$.

Period	*Actual*	*Period*	*Actual*
1	200	5	280
2	220	6	300
3	240	7	320
4	260	8	340

Generalize the procedure which you used so that it can be used to establish initial conditions for problems which exhibit a linear trend.

P11-8. Using the seasonal model, compute a forecast for quarters 2 through 21. Let $A = 0.1$, $B = 0.2$, $\bar{D}_1 = 100$, and the four previous seasonal factors be 1.0, 1.2, 1.0, and .8. Graph your results.

Period	*Actual*	*Period*	*Actual*
1	100	11	101
2	108	12	84
3	99	13	100
4	83	14	142
5	99	15	100
6	128	16	80
7	101	17	99
8	88	18	109
9	100	19	96
10	106	20	71

P11-9. Graph the following time series, determine the most appropriate exponential smoothing model, choose the values of the parameters, prepare a forecast, and compare the results with the actual data.

Period	*Actual*	*Period*	*Actual*
1	80	6	67
2	68	7	61
3	50	8	69
4	66	9	81
5	89	10	71

Period	*Actual*	*Period*	*Actual*
11	51	16	66
12	68	17	85
13	90	18	67
14	73	19	54
15	61	20	65

P11-10. Using the trend, seasonal, and irregular model, compute a forecast for quarters 2 through 25. Assume that the model has been in operation for some time, and that last period's smoothed demand is $\bar{D}_0 = 15$ and its smoothed trend is $\bar{T}_0 = 5$. Let $A = 0.1$, $B = 0.2$, and $C = 0.2$. Finally, let the previous seasonal factors be 1.0, 1.2, 1.0, and 0.8. Graph the forecast and actual.

Period	*Actual*	*Period*	*Actual*
1	21	13	77
2	31	14	104
3	28	15	87
4	27	16	75
5	39	17	96
6	54	18	126
7	49	19	112
8	41	20	95
9	56	21	122
10	79	22	151
11	67	23	134
12	60	24	114

P11-11. Set up a control system to monitor the forecast errors in Prob. 11-10. Let $e_1 = 0.0$ and $\bar{Z}_1 = 0.0$, and let $\overline{\text{MAD}}_1$ be the average of the forecast error for periods 2 through 7. With $K = 2$, is an out-of-control signal sounded? Graph your results.

P11-12. Graph the following time series. Select the most appropriate model. Choose, as best you can, the initial conditions and smoothing constants. Then compute a forecast and compare it graphically with the actual. Finally, develop a control system and graphically depict the behavior of the series.

Period	*Actual*	*Period*	*Actual*
1	34	12	94
2	36	13	116
3	38	14	111
4	47	15	105
5	60	16	120
6	60	17	140
7	60	18	135
8	74	19	126
9	85	20	142
10	85	21	170
11	82		

BIBLIOGRAPHY

Box, G. E. P., and G. M. Jenkins: *Time Series Models for Forecasting and Control,* Holden-Day, San Francisco, 1970.

Brown, Robert G.: *Decision Rules for Inventory Management,* Holt, Rinehart & Winston, Inc., New York, 1967.

———: *Smoothing Forecasting and Prediction of Discrete Time Series,* Prentice-Hall, Inc., Englewood Cliffs, N.J., 1963.

———: *Statistical Forecasting for Inventory Control,* McGraw-Hill Book Company, New York, 1959.

Holt, Charles C., Franco Modigliani, John F. Muth, and Herbert A. Simon: *Planning Production Inventories and Work Force,* Prentice-Hall, Inc., Englewood Cliffs, N.J., 1960.

Winters, Peter R.: "Forecasting Sales by Exponentially Weighting Moving Averages," *Management Science,* April 1960, pp. 324–342.

12

Econometric Forecasting

INTRODUCTION

In the last chapter we studied the problems associated with the forecasting of time series data. The objective was to uncover the way in which a series behaved in the past and extrapolate this pattern into the future.

In this chapter the concern will be with explaining the reasons behind these changes in behavior of a series. Rather than focusing on time as the only explanatory variable, we now choose to determine the underlying causes of these changes. The following case illustrates the central concept in this approach.

Every quarter the production manager of the Generous Skelektric Refrigerator Company determines a sales forecast for the next four quarters. The forecast is used to plan inventory and work force levels for these future periods. As an aid in this forecasting problem, he long ago discovered that certain economic indicators published in government and professional publications provide valuable insight into the way in which refrigerator sales will perform. He discovered, for example, that a forecasted increase in disposable income will usually be followed by an increase in refrigerator sales, and that an increase in housing starts will be followed two quarters later by an increase in refrigerator sales. In addition, there also appeared to be a relationship between price and sales. As the price went up, the volume tended to decrease. The opposite was also true. After observing the relationship between these variables over several quarters, the manager began to rely more and more on them as a source of information from which he compiled his forecast.

His forecasting model, albeit intuitive at this stage, specifies that the sales of refrigerators in any period t can be explained by the forecast of disposable income for that period, the housing starts in period $t - 2$, and the price charged in period t. What the production manager has tried to uncover is the causative factors behind changes in sales volume.

This, then, is the essence of econometric forecasting: the identification of those causative factors which can best explain changes in the variable in which we are interested, and their use for predictive purposes. The difference between econometric forecasting and exponential smoothing should now be clearer. In econometric forecasting we try to uncover *cause-and-effect* relationships, while in exponential smoothing we account for changes in the *time series* by breaking the series into its basic components, including trend, seasonal, cyclical, and irregular. Consequently, a major step in econometric analysis is the identification of variables which might be entered into the analysis. This is not a very straightforward step, for the model builder must select those variables which he feels are the most relevant. At this stage only experience and intuition can help.

In this chapter several econometric models will be developed. In the first model a simple linear relationship is assumed to exist between two variables. Next, models are explored where the relationships are nonlinear; and then a model is developed where several causative factors may be operating: this is the multivariate case. Finally, the theory of econometrics is explored, and we consider the assumptions of the models together with the consequence of their violation.

SIMPLE LINEAR REGRESSION

In the two-variable case we have a variable whose changes we would like to predict and a variable which it is hoped can explain these changes. For example, refrigerator sales might be the variable which is to be predicted and new housing starts might be the explanatory variable. The variable which we would like to predict is called the *dependent* variable, while the explanatory variable is called the *independent* variable. Denoting the dependent variable by Y and the independent variable by X, we have

$$Y = f(X)$$

or Y is a *function* of X. This is merely another way of stating that X influences Y.

THE FORM OF THE RELATIONSHIP

The simplest relationship between Y and X is a linear one:

$$Y = a + bX$$

This is the equation for a straight line where a is the intercept and b the slope. Consider the relationship shown in Fig. 12-1. The straight line intercepts the Y axis at point a and has a slope which can be expressed as a change in Y for a given change in X, or $\Delta Y/\Delta X$. Therefore slope $= b = \Delta Y/\Delta X$. Furthermore, the slope of a straight line never changes.

Consider the following example:

Refrigerator sales (millions of \$) Y	*Housing starts (thousands of units)* X
2	10
4	30
3	20
5	40

If a plot is made of these two variables, a straight line drawn through these points gives a perfect fit. In fact, it can be seen from Fig. 12-2 that the equation of the line can be estimated as

$$Y = 1 + .10X$$

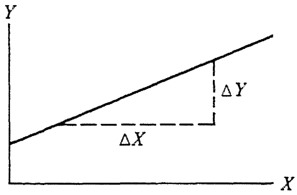


Fig. 12-1 Slope of a line.

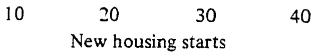


Fig. 12-2 Simple linear regression relationship.

Since all the points fall on the line, we can say that the relationship between these variables is a perfect one.

The purpose of determining this relationship is to use it for predictive purposes. Therefore, if the level of housing starts for next period was estimated to be 25, the estimated sales for the company would be

$$Y = 1 + .10X$$
$$Y = 1 + .10(25)$$
$$Y = 2.5$$

Since the relationship as evidenced in the plot is a strong one, a lot of confidence could be placed in this estimate.

Consider, on the other hand, the following set of data.

Refrigerator sales Y	*Housing starts* X
2.2	10
4.2	30
2.5	20
4.5	40

It can be seen in Fig. 12-3 that the points do not lie along a straight line; therefore the *degree* of relationship is not nearly as strong as in the first case. It appears, however, that the *direction* of the relationship, which is given by the equation of the line, is the same:

$$Y = 1 + .10X$$

We have just touched upon the two major aspects of econometric forecasting. These include regression and correlation analysis. In regression analysis the direction of the relationship is determined. Rather than fairing in a line by eye, we will shortly develop a statistical methodology for obtaining a fit. Correlation analysis, on the other hand, concerns the degree of the relation-

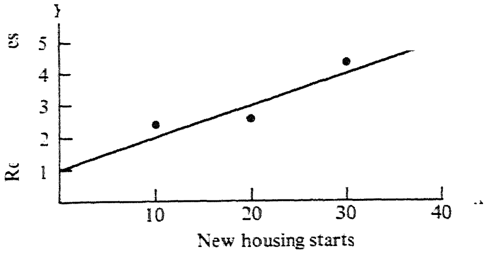


Fig. 12-3 Simple linear regression relationship.

ship. In Fig. 12-4 different degrees of relationship are shown. In (*a*) the degree of relationship is perfect, and we say the correlation is 1. In (*b*) there is apparently no relationship between the dependent and independent variables and their correlation is zero. In (*c*) we have a situation in between these two extremes. The correlation here would therefore be between 0 and 1.

Linear relationships are certainly not the only possible ones. Later in the chapter nonlinear relationships will be considered, but first we will develop the simple two-variable linear case.

DISTURBANCE TERM

If we specify the regression relationship between Y and X in the following way:

$$Y = a + bX$$

this implies that, given a value of X, we know the *exact* value of Y. In economic analysis this is never the case.

Let us reconsider the example of refrigerator sales and housing starts. If many observations at *each* level of housing starts could be taken, we would hardly expect to observe the same sales figure for each particular level. In some years the sales of refrigerators for this given level of housing starts would be different from the sales in other years. This idea, for several levels of housing starts, is expressed in Fig. 12-5. At level X_1 five observations on Y are seen; six at level X_2; and so on.

Unknown to us, a relationship between X and Y exists which has generated these observations. This we will call the *population regression equation* and

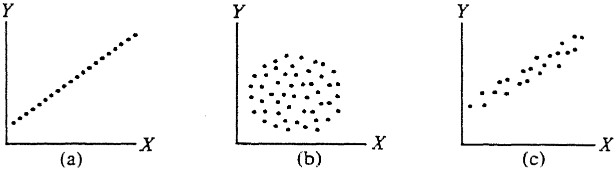


Fig. 12-4 Examples of correlation.

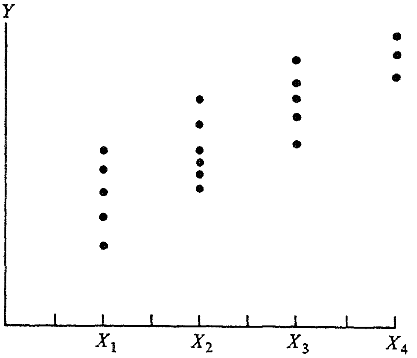


Fig. 12-5 Observed variability in Y.

express it as

$$Y_i = a + bX_i + \mu_i$$

where μ_i = disturbance term
a, b = unknown population parameters we would like to estimate

Considering one set of observations at X_1, their mean value will be at $a + bX_1$, but actual observations can be specified by

$$Y_1 = a + bX_1 + \mu_1$$
$$Y_2 = a + bX_1 + \mu_2$$
etc.

The values of $\mu_1, \mu_2, \ldots$, indicate the amount by which the sales for a particular level of housing starts fall short of or exceed the central value $a + bX_1$.

Perhaps the main justification for including this stochastic or disturbance term is that it is not possible to identify every source of the variation in Y. Here we have selected X, but surely other unidentified sources still exist. Some of these other sources might include the price that the firm charges for its refrigerators, the price of competitors' refrigerators, the disposable income of individuals, the number of new household formations, consumer expectations for future income levels, Paris fashions, the price of sports cars, and so on. The list goes on and on, including both quantifiable and nonquantifiable causative factors. It is unreasonable to expect to identify and obtain data on all these variables. We therefore attempt to identify the important explanatory variable(s) and let the net effect of the excluded variables be incorporated in the error term μ_i. This term then represents the amount of variation in the dependent variable that can be attributed to unassignable causes. Since many of these unidentified factors are operating at any one time, it is reasonable to assume that many of these will be pulling in opposite directions. It is very unlikely that they will all pull in the same direction. Therefore small values of the disturbance term μ_i are much more likely than large ones.

There are two other ways in which we can justify the inclusion of the error term. One is that even if all identifiable factors were considered, there is still a basic unpredictable randomness in nature. Finally, there is measurement error. In measuring the level of any economic activity, it is very difficult to determine the true level of that activity.

In summary, then, there is certainly reason enough to include the error term, and we now can write the theoretical regression relationship of the population as

$$Y_i = a + bX_i + \mu_i$$

ASSUMPTIONS OF THE REGRESSION MODEL

Some assumptions must be made about this disturbance term before we can proceed with our analysis. First, the mean of the error term distribution, for

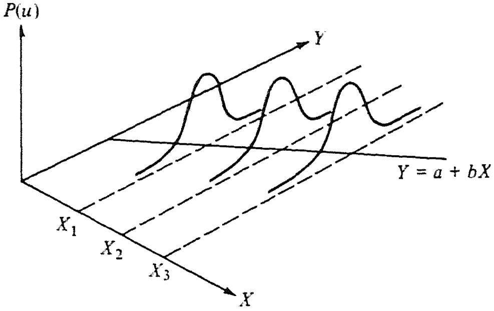


Fig. 12-6 Error term distribution.

any value of the independent variable, must be zero. This can be seen in Fig. 12-6. The vertical axis, $P(\mu)$, represents the probability of different values of the error term. According to this first assumption, this distribution is centered on the regression line. It can also be seen from this plot that smaller values of μ are much more likely than larger ones.

The second assumption relates to the variability of the disturbance term: it must be constant. Therefore the variability of the disturbance term shown at X_1 in Fig. 12-6 must be the same as that at $X_2, X_3, \ldots X_n$.

The final assumption is that the values of the disturbance term must be independent of one another. For example a large positive disturbance term observed at one level of X_1 should not be consistently followed by a large positive disturbance term when an observation is next made at X_2.

In economic analysis these assumptions are quite important, and later in the chapter we will examine the way in which the analyst can determine whether these assumptions have been violated and the consequences of such an occurrence.

LINE OF LEAST SQUARES

Assume that a sample of n observations has been taken on refrigerator sales Y and housing starts X. A plot of these data in Fig. 12-7 indicates that a strong relationship exists between them. Our first step in this analysis will be to estimate the equation for the line which best fits these data. Now, the *true* population relationship

$$Y_i = a + bX_i + \mu_i$$

is unknown to us. We can, however, estimate this relationship from our sample observations. Since we compute but an estimate, our statistical analysis will provide us with

$$\hat{Y}_i = \hat{a} + \hat{b}X_i$$

where $\hat{a}, \hat{b}$ = estimates of the two unknown parameters a and b
$\hat{Y}_i$ = estimated level of activity for any value of X_i

First a criterion for determining $\hat{a}$ and $\hat{b}$ must be chosen. Perhaps the best-known is that of "least squares." Returning to Fig. 12-7, if we take any

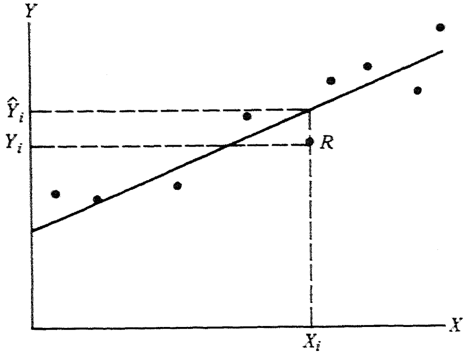


Fig. 12-7 Line of least-squares example.

point R with coordinates Y_i, X_i, the actual level of the dependent variable is Y_i and its estimated value is $\hat{Y}_i$. Drawing a perpendicular line from R, the error is

$$e_i = \hat{Y}_i - Y_i$$

The errors will be positive or negative for each observation of X. If we square the error for all observations, the sum will give some indication of the spread of these points about the line. The principle of least squares is to choose those values of $\hat{a}$ and $\hat{b}$ which will minimize the sum of these error terms squared:

$$\text{Minimize} \sum_{i=1}^{n} e_i^2$$

Substituting for e_i we have

$$\text{Minimize} \sum_{i=1}^{n} (\hat{Y}_i - Y_i)^2$$

or

$$\text{Minimize} \sum_{i=1}^{n} (\hat{a} + \hat{b}X_i - Y_i)^2$$

Using calculus, the following set of *normal equations* can be derived:

$$\sum_{i=1}^{n} Y_i = n\hat{a} + \hat{b} \sum_{i=1}^{n} X_i$$

$$\sum_{i=1}^{n} X_iY_i = \hat{a} \sum_{i=1}^{n} X_i + \hat{b} \sum_{i=1}^{n} X_i^2$$

where n = number of observations

From these equations and the data on X_i and Y_i, values of $\hat{a}$ and $\hat{b}$ can be computed such that the sum of the error terms squared will be less for these values than for any others. Now we turn to an example.

Table 12-1 Computations necessary for solving normal equations

Refrigerator sales Y	*Housing starts* X	XY	X^2
2.2	10	22	100
4.2	30	126	900
2.5	20	50	400
4.5	40	180	1,600
$\Sigma Y = 13.4$	$\Sigma X = 100$	$\Sigma XY = 378$	$\Sigma X^2 = 3{,}000$

Example Consider the data given in Fig. 12-3. From the normal equations it can be seen that we must determine ΣY, ΣX, ΣXY, and ΣX^2. These computations are performed in Table 12-1.

Substituting into the normal equations

$$\begin{aligned} 13.4 &= 4\hat{a} + 100\hat{b} \\ 378 &= 100\hat{a} + 3{,}000\hat{b} \end{aligned}$$

and solving simultaneously by multiplying the first equation by 25 and subtracting the second equation from the first, we have

$$\begin{array}{rcr} 335 &=& 100\hat{a} + 2{,}500\hat{b} \\ -378 &=& -100\hat{a} - 3{,}000\hat{b} \\ \hline -43 &=& -500\hat{b} \\ \hat{b} &=& .086 \end{array}$$

Finally we substitute b into the first equation and solve for a.

$$\begin{aligned} 13.4 &= 4\hat{a} + 100(.086) \\ \hat{a} &= \frac{13.4 - 8.6}{4} \\ \hat{a} &= 1.2 \end{aligned}$$

Therefore the regression equation can be written as

$$\hat{Y} = 1.2 + .086X$$

Next we turn to the problem of determining the degree of this relationship.

COEFFICIENT OF DETERMINATION

Instead of both a dependent and an independent variable from which to draw a relationship, consider having data on just the dependent variable Y. In our example we would have data on only the level of refrigerator sales for four periods. With nothing else available, the best prediction would be based on

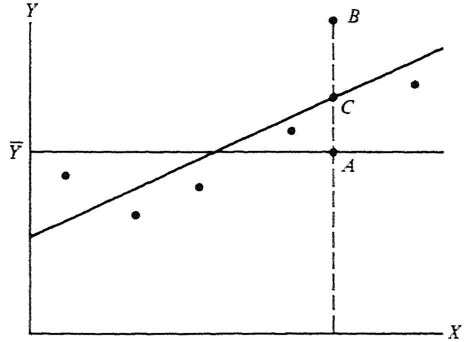


Fig. 12-8 Error explained by the introduction of the linear relationship.

a simple average of these data. By calculating the sum of the squared deviations from the mean,

$$\text{SST} = \sum_{i=1}^{n} (Y_i - \bar{Y})^2$$

we have a measure of the "sum of squares total." This is the magnitude of the error when the mean of the Y_i values is used to represent the average level of sales.

Now we introduce the independent variable X and analyze the effect it has on this "sum of squares total." In Fig. 12-8 the average value for the independent variable $\bar{Y}$ is shown. If we look at point B, we can see that before the linear relationship was introduced, the error of estimation was AB; once the linear relationship was introduced, the error was reduced to BC. Therefore AC represents the amount that was explained by introducing the linear relationship. Measuring the *total* amount of error that has been explained away by the introduction of the regression line, we have

$$\text{SSR} = \sum_{i=1}^{n} (\hat{Y}_i - \bar{Y})^2$$

where SSR = sum of squares explained by regression

We can now determine the effect of introducing the linear trend by the ratio

$$\frac{\text{SSR}}{\text{SST}} = \frac{\Sigma(\hat{Y}_i - \bar{Y})^2}{\Sigma(Y_i - \bar{Y})^2}$$

This represents the proportion of the Y variation explained by the linear influence of X. This ratio is called the *coefficient of determination* r^2. The *coefficient of correlation* is the square root of this number, or r. The meaning behind the coefficient of determination is much more intuitive than that behind the coefficient of correlation. It is the percent of the variability that is explained by the linear relationship. Therefore an r^2 of .9 indicates that 90 percent of the variability has been explained by the introduction of the linear relationship.

Table 12-2 Computation of coefficient of determination

Y_i *Refrigerator sales*	X_i *Housing starts*	$\hat{Y}_i$	$(\hat{Y}_i - \bar{Y})$	$(\hat{Y}_i - \bar{Y})^2$	$(Y_i - \bar{Y})$	$(Y_i - \bar{Y})^2$
2.2	10	2.06	−1.29	1.66	−1.15	1.32
4.2	30	3.78	.43	.18	.85	.72
2.5	20	2.92	− .43	.18	− .85	.72
4.5	40	4.64	1.29	1.66	1.15	1.32
				3.68		4.08

Example Now we return to our example and calculate the coefficient of determination. The computations are shown in Table 12-2.

$$r^2 = \frac{\text{SSR}}{\text{SST}} = \frac{3.68}{4.08} = .92$$

It can therefore be concluded that 92 percent of the variability associated with refrigerator sales has been explained by the introduction of the variable housing starts.

NONLINEAR RELATIONSHIPS

In the preceding section linear relationships were considered. There are of course many times when relationships display nonlinear characteristics. Consider the plot in Fig. 12-9. If a straight line of the form

$$\hat{Y} = \hat{a} + \hat{b}X$$

were fit to the data, we would obtain

$$\hat{Y} = -0.3 + 0.2X$$

with a coefficient of determination of

$$r^2 = .95$$

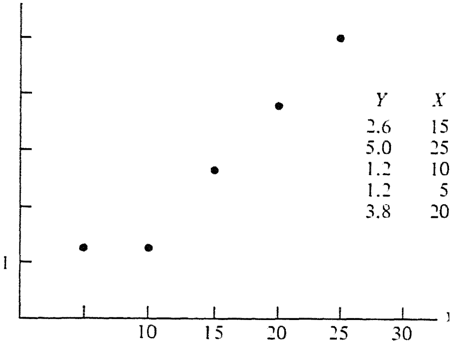


Fig. 12-9 Nonlinear regression example.

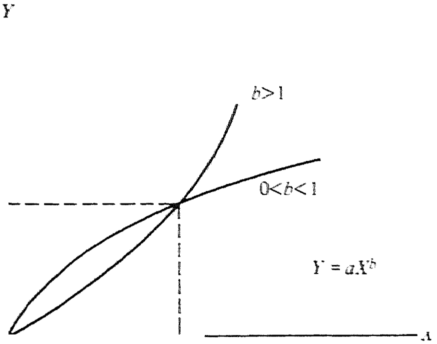


Fig. 12-10 Power function $Y = aX^b$.

Upon closer inspection it is apparent that a nonlinear function might prove to be a better fit.

Of the many types of nonlinear relationships, two will be considered here. These two are chosen because they represent two of the more common relationships that are found in economic data. They include the power and parabolic functions.

POWER FUNCTION

A power function takes the following form:

$$Y = aX^b$$

The letter b represents the slope of the function, and, if b is greater than 1, the *slope continually increases*. On the other hand, if b is between zero and 1, the slope continually decreases. This can be seen in Fig. 12-10. Another property of this family of curves is that they all pass through a,1.

If the function is transformed into logs, it becomes

$$\log Y = \log a + b \log X$$

Furthermore, if we let

$$Y' = \log Y$$

$$a' = \log a$$

$$X' = \log X$$

we then have

$$Y' = a_1 + bX'$$

and can conclude that the transformed function is linear.

Therefore when it is suspected that the power function would fit the data well, the logs of Y, a, and X can be found. The logs are then used in the normal equations to find the least-squares estimates of a' and b. Finally, the antilog must be taken of a' to get a; then we can write

$$Y = aX^b$$

The coefficient of determination can be calculated just as we did in the linear case. If the nonlinear function is a better fit, its value of r^2 will be higher.

PARABOLIC FUNCTIONS

The second set of relationships can be expressed as parabolic functions of the following type:

$$Y = a + bX + cX^2$$

This particular one is referred to as a *second-degree parabola,* where the degree is determined by the largest exponent. The form of these curves is illustrated in Figs. 12-11 and 12-12. In Fig. 12-11 a family of parabolas is represented whose value of c is less than zero, while in Fig. 12-12 the case where c is greater than zero is presented.

It is important to recognize that, since the parabola is an arc, there is usually no logical basis for extrapolation. For example, we might determine from a data plot that it can best be fit by the first half of the parabola shown in Fig. 12-11. To use this relationship for predictive purposes when the independent variable is greater than $-b/2c$ would be very questionable.

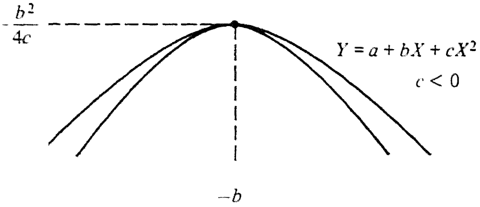


Fig. 12-11 Parabola where $c < 0$.

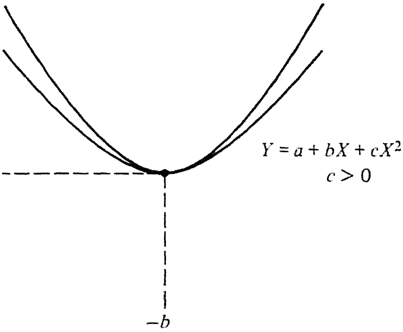


Fig. 12-12 Parabola where $c > 0$.

Returning to the example illustrated in Fig. 12-9, it appears that a parabola of the form

$$Y = a + bX + cX^2$$

would be a better fit than a linear one. The power function is probably not appropriate, since that function,

$$Y = aX^b$$

must pass through the origin, $X = 0$ and $Y = 0$.

If we apply the least-squares principle to the equation for a second-degree parabola, we get the following set of normal equations:

$$\begin{aligned} \Sigma Y_i &= na + b\Sigma X_i + c\Sigma X_i^2 \\ \Sigma X_i Y_i &= a\Sigma X_i + b\Sigma X_i^2 + c\Sigma X_i^3 \\ \Sigma X_i^2 Y_i &= a\Sigma X_i^2 + b\Sigma X_i^3 + c\Sigma X_i^4 \end{aligned}$$

The fit obtained by using these equations can be measured by r^2 just as we did in the nonlinear case.

If we use these normal equations to determine the regression equation in Fig. 12-9, we get

$$\hat{Y}_i = .8 + 0.015X + 0.006X^2$$

and $r^2 = .98$. Since the r^2 generated by the nonlinear analysis is higher than that from the linear analysis, the former is a better fit.

Perhaps the simplest way to determine whether the linear relationship would be adequate is to draw a scatter plot of the data. From the plot and some familiarity with these and other standard nonlinear functions, the analyst will have a good idea of where to start.

MULTIVARIATE REGRESSION ANALYSIS

We now turn to the more realistic case where variations in the dependent variable are related to more than one independent variable. For example, it might seem reasonable to expect that the level of sales for refrigerators can be better explained by both housing starts and disposable income. Mathematically we can express this *multivariate* relationship in general form as

$$Y = f(X_2, X_3)$$

where Y = level of refrigerator sales

X_2 = level of housing starts

X_3 = level of disposable income

If it is reasonable to assume that the relationship is linear, the explicit form of the population regression equation can be written as

$$Y_i = A + BX_{2i} + CX_{3i} + \mu_i$$

The *estimate* of this relationship can be expressed as

$$\hat{Y}_i = \hat{a}_{1.23} + \hat{b}_{12.3}X_{2,i} + \hat{b}_{13.2}X_{3,i}$$

The equation now represents a plane which is fitted to a set of data points in three-dimensional space (Y, X_2, and X_3).

The constant $a_{1.23}$ is the intercept formed by the regression plane on the Y axis and is the estimate of A. In subscripting the constant, the number before the decimal point represents variable 1 or the dependent variable. The numbers after the decimal point represent the independent variables included in the regression analysis. The coefficient of the first variable $\hat{b}_{12.3}$ is the estimate of B and denotes the amount by which a unit change in X_2 is expected to affect Y. The second coefficient $\hat{b}_{13.2}$ is an estimate of C and denotes the way in which a unit change X_3 affects Y. The reason for this elaborate system of subscripts is that it clearly indicates the number of variables in the analysis. For example, in the subscript associated with $\hat{b}_{12.3}$, the first number denotes that the variable on the left-hand side of the equation, or the dependent variable, is the variable 1. The second number indicates the variable to which the coefficient is attached, and the number (or numbers) after the decimal point represent other variables that have been taken into consideration in the estimation of this regression equation. For example, if a relationship is hypothesized between four variables, the estimate of the relationship can be expressed as

$$\hat{Y}_i = a_{1.234} + b_{12.34}X_{2i} + b_{13.24}X_{3i} + b_{14.23}X_{4i}$$

METHOD OF LEAST SQUARES

Again we utilize the least-squares criterion in developing the normal equations. For the three-variable case, estimates for the constant term and two coefficients can be found by solving the following normal equations:

$$\begin{aligned}
\Sigma Y &= \hat{a}_{1.23}n + \hat{b}_{12.3}\Sigma X_2 + \hat{b}_{13.2}\Sigma X_3 \\
\Sigma X_2 Y &= \hat{a}_{1.23}\Sigma X_2 + \hat{b}_{12.3}\Sigma X_2^2 + \hat{b}_{13.2}\Sigma X_2 X_3 \\
\Sigma X_3 Y &= \hat{a}_{1.23}\Sigma X_3 + \hat{b}_{12.3}\Sigma X_2 X_3 + \hat{b}_{13.2}\Sigma X_3^2
\end{aligned}$$

The job of solving three simultaneous equations is laborious, but nonetheless it is possible to solve the problem by hand. In the case of four variables, the appropriate set of normal equations would include four equations, and it is clear that the process of solving this and larger problems would be getting out of hand. Fortunately most computer centers maintain a regression code that the uninitiated can become familiar with in a few minutes' time. For problems of a realistic scale, then, a computer code is a must.

Example Now let us turn to an example. The data are given in Table 12-3, along with the necessary mathematics for computing the figures needed

Table 12-3 Computations necessary for solving normal equations

Observation	Y	X_2	X_3	X_2Y	X_3Y	X_2X_3	X_2^2	X_3^2
1	2	3	2	6	4	6	9	4
2	5	6	3	30	15	18	36	9
3	3	3	2	9	6	6	9	4
4	2	3	2	6	4	6	9	4
5	2	2	3	4	6	6	4	9
6	3	4	1	12	3	4	16	1
7	5	6	4	30	20	24	36	16
8	3	3	4	9	12	12	9	16
9	1	2	0	2	0	0	4	0
10	2	3	1	6	2	3	9	1
	28	35	22	114	72	85	141	64
	ΣY	ΣX_2	ΣX_3	ΣX_2Y	ΣX_3Y	ΣX_2X_3	ΣX_2^2	ΣX_3^2

in the normal equations. The normal equations can then be set up as follows:

$$\begin{aligned} 28 &= 10\hat{a}_{1.23} + 35\hat{b}_{12.3} + 22\hat{b}_{13.2} \\ 114 &= 35\hat{a}_{1.23} + 141\hat{b}_{12.3} + 85\hat{b}_{13.2} \\ 72 &= 22\hat{a}_{1.23} + 85\hat{b}_{12.3} + 64\hat{b}_{13.2} \end{aligned}$$

Solving simultaneously, we get

$$\begin{aligned} \hat{a}_{1.23} &= -.42 \\ \hat{b}_{12.3} &= .74 \\ \hat{b}_{13.2} &= .29 \end{aligned}$$

Therefore the regression equation can be written as

$$\hat{Y}_i = -.42 + .74X_{2i} + .29X_{3i}$$

COEFFICIENT OF MULTIPLE DETERMINATION

We have already seen that the coefficient of determination for the simple two-variable case is defined as the explained variability as a percentage of the total variability. The *coefficient of multiple determination* for functions of three or more variables is defined in a similar way.

$$R = \frac{\text{SSR}}{\text{SST}}$$

where R = coefficient of multiple determination

SSR = sum-of-squares explained by regression

SST = sum-of-squares total

Table 12-4 Computation of coefficient of determination

Y_i	SST $(Y_i - \bar{Y})^2$	$\hat{Y}_i$	SSR $(Y_i - \bar{Y})^2$
2	$(-.8)^2 =$ .64	2.4	$(-.4)^2 =$ .16
5	$(2.2)^2 =$ 4.84	4.9	$(2.1)^2 =$ 4.41
3	$(.2)^2 =$ .04	2.4	$(-.4)^2 =$ .16
2	$(-.8)^2 =$ .64	2.4	$(-.4)^2 =$ .16
2	$(-.8)^2 =$ .64	1.9	$(-.9)^2 =$ .81
3	$(.2)^2 =$ .04	2.8	$(0)^2 =$ 0.0
5	$(2.2)^2 =$ 4.84	5.2	$(2.4)^2 =$ 5.76
3	$(.2)^2 =$ .04	2.9	$(.1)^2 =$ .01
1	$(-1.8)^2 =$ 3.24	1.1	$(-1.7)^2 =$ 2.89
2	$(-.8)^2 =$ .64	2.1	$(-.7)^2 =$ .49
	15.60		14.85

The computations for the problem presented in Table 12-3 are carried out in Table 12-4 and the coefficient of determination is computed to be

$$R = \frac{14.85}{15.6} = .95$$

//// JOINT TEST FOR THE SIGNIFICANCE OF ALL REGRESSION COEFFICIENTS

The n observations that were used in determining an estimate of the regression equation represent but a sample set of observations from the true but unknown population

$$Y_i = A + BX_{2i} + CX_{3i} + u_i$$

If another sample of n observations were taken, another estimate of the regression equation would be obtained with a different constant and different regression coefficients. We can conclude that the set of coefficients that we get in a regression analysis is but an estimate. It might therefore be useful to determine whether these coefficients differ significantly from zero. If they do not, the independent variables contribute nothing to the explanation of changes in Y, and the coefficient of determination as well as the coefficients in the regression equation were but chance variations from zero. It is beyond the scope of this book to look at this analysis in detail. The student is therefore referred to a book on statistics or econometrics for a full, theoretical development. While sacrificing the theoretical detail we will, however, turn to the practical aspects involved in testing for a significant difference from zero.

First we must define an F ratio as

$$F = \frac{\text{SSR}/(k-1)}{\text{SSE}/(n-k)}$$

where SSR = sum of squares explained by regression, $\Sigma(\hat{Y}_i - \bar{Y})^2$
SSE = sum of squares due to error, or SSE = SST − SSR
$\text{SSE} = \Sigma(Y_i - \hat{Y}_i)^2$
k = number of variables, both dependent and independent, in regression relationship
n = number of observations

Now, this ratio has an F distribution with $k - 1$ and $n - k$ *degrees of freedom.* At the expense of oversimplification, the degrees of freedom are the number of observations, n, in the sample minus the number of parameters which must be estimated by sample observations. Therefore in the sum-of-squares error, the number of degrees of freedom is $n - k$, since there are $k - 1$ coefficients of independent variables and one constant term, for a total of k parameters which must be estimated. The number of degrees of freedom associated with the sum-of-squares total is $n - 1$, since the mean is the only parameter that must be estimated.

A property of these sums of squares is that the sum-of-squares total is equal to the sum-of-squares regression plus the sum-of-squares error.

$$\text{SST} = \text{SSR} + \text{SSE}$$

This also holds for the degrees of freedom df, where

$$df_{\text{SST}} = df_{\text{SSR}} + df_{SSE}$$

Therefore, to determine the number of degrees of freedom associated with the sum-of-squares regression, we might take the number of degrees of freedom associated with the sum-of-squares total and subtract from it the number of degrees of freedom associated with the sum-of-squares error:

$$df_{\text{SSR}} = df_{\text{SST}} - df_{\text{SSE}}$$

$$df_{\text{SSR}} = (n - 1) - (n - k)$$

$$df_{\text{SSR}} = k - 1$$

The number of degrees of freedom associated with the sum-of-squares regression is therefore $k - 1$.

The F distribution is a mathematical distribution for which a table is found in Appendix B. The table is entered in the top row with the number of degrees of freedom associated with the sum-of-squares regression $k - 1$. Down the first column the table is entered with the number of degrees of freedom associated with the sum-of-squares error $n - k$. The critical F ratio can be found where these intersect. To illustrate the way in which we read the table, consider the present example. The degrees of freedom for the sum-of-squares regression is $k - 1 = 2$ and the degrees of freedom for the sum-of-squares error is $n - k = 7$. Entering the table for 2 and 7 degrees of freedom, the critical value of F is found to be 4.74 at the 5 percent level and 9.55 at the 1 percent level. First we will interpret the figure at the 5 percent level. There is only a 5 percent chance that a computed F value greater than 4.74 can occur when in fact the coefficients are not significantly different from zero.

In our example, the computed F value is

$$F = \frac{\text{SSR}/(k-1)}{\text{SSE}/(n-k)} = \frac{14.85/2}{.75/7} = \frac{7.42}{.11} = 68$$

where SSE = 15.60 — 14.85 = .75.

Only 5 percent of the time will the computed F value be larger than 4.74 when in fact the true but unknown regression coefficients are not significantly different from zero. Since 68 is much larger than 4.74, we can conclude that the regression results are significant at the 5 percent level. That is, more than just a chance variation from zero explains the magnitude of the regression coefficients.

Looking at the critical F value for 1 percent, we can see that only 1 percent of the time will the computed F value be larger than 9.55 when the coefficients are not significantly different from zero. Since our computed F value is greater than the critical F value, the results are also significant at the 1 percent level.

The test just developed determines the significance of the regression equation. If the conclusion is that the coefficients, when tested jointly by the F test, are not significant, the model builder must abandon his present hypothesis and search for others. There is, however, another test that allows one to test the significance of each coefficient singly rather than testing jointly. In this way those variables which can significantly explain the variability in the dependent variable can be used in developing a new hypothesis.[1] ////

THE THEORY OF ECONOMETRICS

The validity of the F test and other statistical tests of the regression equation depends upon whether or not the assumptions made for the disturbance term have been violated. These assumptions included a normally distributed error term with constant variance and serial independence. When the assumptions are violated, the conclusions derived from these standard tests may be invalid. As you will shortly see, it is often possible to overcome these problems and improve predictive ability.

NORMAL ERROR TERM

The first assumption specifies that the error term must be normally distributed about the regression line with mean zero. Therefore large error excursions are less likely than small ones, and the greatest concentration will be found around zero. We concluded earlier in this chapter that the error term can be largely attributed to the omission of a host of independent variables from the analysis. It turns out that small deviations from normality will not adversely affect these tests. Violations of the other assumptions are, however, more serious.

---

[1] J. Johnston, *Econometric Methods*, McGraw-Hill Book Company, New York, 1963, p. 118.

HOMOSCEDASTICITY

The second assumption is that the error term must have a constant variance over all values of the independent variable. Consider the simple regression relationship between refrigerator sales and housing starts. The assumption of homoscedasticity specifies that if many observations were taken at each level of housing starts, the variance at each level would be the same. If the variance, for example, increased with a greater number of housing starts, we would have the case of a heteroscedastic error term. Figure 12-13*a* represents the homoscedastic case while Fig. 12-13*b* represents the heteroscedastic case.

It is usually difficult to test for heteroscedasticity. A test such as this would call for a large number of observations at each level of the independent variable, and in most cases only one observation at each level is available. In the absence of a perfect test, several levels of housing starts could be grouped together. The standard deviation of the observations for each group could be compared with that of each other group to see that their size does not increase significantly. If the variability does increase, the remedy is usually to transform the data into logarithms before the regression correlation analysis is begun.

AUTOCORRELATION

The problem of autocorrelation occurs when successive error terms are not independent of one another. Consider the case where time is the independent variable and refrigerator sales the dependent variable. It can be seen from Fig. 12-14 that the regression line fits the average trend well; however, there appears to be a seasonal factor present. The error terms from one period to the next are clearly not independent. The magnitude of the error term in one instant of time depends upon the magnitude of the error in the previous period. We therefore conclude that the error terms are autocorrelated.

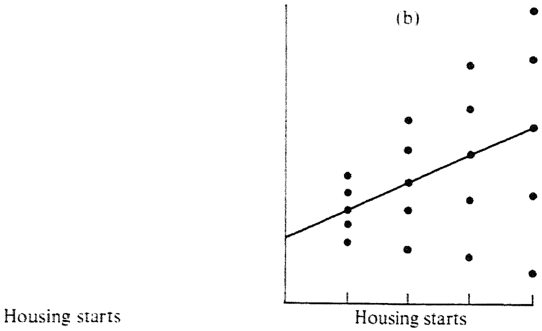


Fig. 12-13 (*a*) Homoscedasticity; (*b*) Heteroscedasticity.

Fig. 12-14 Autocorrelation.

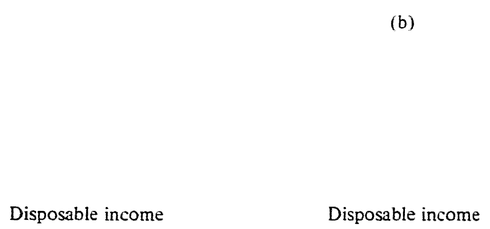


Fig. 12-15 Elimination of autocorrelation.

Consider another example of autocorrelation illustrated in Fig. 12-15. The dependent variable is the sales of TV sets and the independent variable is disposable income. A linear relationship is initially hypothesized and the regression line is shown in Fig. 12-15*a*. The error terms around this regression line are certainly not independent. This autocorrelated error term can be attributed to the improper specification of the form of the regression line. Rather than a linear relationship, a nonlinear relationship of the type shown in Fig. 15-12*b* should have been specified.

A common test for autocorrelation is to compute the Durbin/Watson statistic. In fact this statistic is provided along with the regression results of many good computer codes used for regression and correlation analysis.

The consequence of autocorrelation is that tests of significance, such as the *F* test, may be seriously affected. The solution to the problem, then, is to remove the autocorrelation. In the case of the improper specification of the regression line, additional variables might be included so that a nonlinear function could be fitted to the data. Another set of methods for removing autocorrelation involves the transformation of data.[2]

MULTICOLLINEARITY

In multiple regression analysis it often happens that there is a strong relationship between the independent variables themselves. Consider a relationship between sales of refrigerators as the dependent variable with disposable income and housing starts as the independent variables. It might be found that a strong correlation exists between the two independent variables. That is, if disposable income was plotted against housing starts, the scatter of points would cluster closely around the regression line. Let us assume the correlation was perfect. Under this condition it would be impossible to measure the separate influences that each independent variable had upon the dependent variable. There is no way to determine how housing starts by itself would affect sales with disposable income held constant. Therefore, if perfect correlation is present, either variable could be used in place of the other. In the usual case where multicollinearity is present, the degree of correlation between independent variables is not perfect but high. In this situation it may not be possible to find the regression coefficients with any degree of accuracy. Therefore in the presence of multicollinearity the variability associated with these independent variables can be quite high.

---

[2] Michael J. Brennan, *Preface to Econometrics,* South-Western Publishing Company, Inc., Cincinnati, 1965, p. 359.

To avoid the problem of multicollinearity, variables should be chosen which have little correlation between themselves. In some computer programs a correlation matrix is printed out. From this one can determine the magnitude of the multicollinearity problem by observing the degree of correlation that exists between the independent variables included in the regression equation.

DUMMY VARIABLES

In some cases the model builder is faced with incorporating temporal and nonquantitative explanatory variables into the regression analysis. For example one might initially hypothesize that the sales of refrigerators is linearly related to disposable income. But upon closer analysis of the periods under study it is found that during the wage and price freeze of 1971 the sales function temporarily shifted upward. Rather than discarding these data points, two dummy variables X_2 and X_3 might be included such that

$$X_2 = \begin{cases} 1 & \text{during the freeze} \\ 0 & \text{otherwise} \end{cases}$$

$$X_3 = \begin{cases} 1 & \text{otherwise} \\ 0 & \text{during the freeze} \end{cases}$$

Our hypothesis, then, can be expressed in the following way:

$$Y_i = BX_{2i} + CX_{3i} + DX_{4i} + \mu_i$$

where Y_i = refrigerator sales

$$X_{2i} = \begin{cases} 1 & \text{during the freeze} \\ 0 & \text{otherwise} \end{cases}$$

$$X_{3i} = \begin{cases} 1 & \text{otherwise} \\ 0 & \text{during the freeze} \end{cases}$$

X_{4i} = disposable income

We have implied that the coefficient for X_{2i} is the intercept for the freeze period and that the coefficient of X_{3i} is the intercept for nonfreeze periods.

There is, however, a serious difficulty in this formulation. The regression equation does not have a constant term. As such it could not be solved. To overcome this problem, we simply let one of the dummy variables be zero. We now have

$$Y_i = A + BX_{2i} + CX_{3i} + \mu_i$$

$$\text{where } X_{2i} = \begin{cases} 0 & \text{during freeze period} \\ 1 & \text{otherwise} \end{cases}$$

This gives a freeze-period sales function as

$$Y_i = A + CX_{3i} + \mu_i$$

and a nonfreeze sales function as

$$Y_i = (A + B) + CX_{3i} + \mu_i$$

As illustrated in the last example, dummy variables can be used when

1. The observations can be divided into mutually exclusive categories, and
2. The effect of the categories is to shift the regression equation (change the intercept) without affecting the slope (as in all our examples), or to shift the coefficient of a variable.[3]

It was also pointed out that the regression equation must be so structured as to ensure a constant term.

Of the many uses for dummy variables another is for seasonal changes. For example, we could define the following variables

$$X_1 = \begin{cases} 1 & \text{in the fall} \\ 0 & \text{otherwise} \end{cases}$$

$$X_2 = \begin{cases} 1 & \text{in the winter} \\ 0 & \text{otherwise} \end{cases}$$

$$X_3 = \begin{cases} 1 & \text{in the spring} \\ 0 & \text{otherwise} \end{cases}$$

Dummy variables can also be used to represent other qualitative variables such as sex, marital status, education, occupation, and social status.[4] Now we turn to an example.

Example The marketing manager of a large electronics firm feels that his sales of TV sets is quite responsive to the advertising budget. He would like to determine as closely as possible the nature of this relationship (that is, the direction and degree).

Muddying the water, however, is the fact that the fall-winter season invariably generates a higher level of sales than does the spring-summer season. He feels that this factor should also be considered in determining the basic relationship. Finally he decided to confront his statistical department with the data presented in Table 12-5. They formulate the following hypothesis:

---

[3] J. Johnston, op. cit., p. 221.

[4] J. Johnston, op. cit., p. 221.

$$Y_i = A + BX_{2i} + CX_{3i} + \mu_i$$

where Y_i = TV sales over 6-month period i

$$X_{2i} = \begin{cases} 1 & \text{in fall-winter sales period} \\ 0 & \text{otherwise} \end{cases}$$

X_{3i} = advertising budget over 6-month period i

The regression equation is estimated to be

$$Y = 4.47 + 25.6X_2 + .275X_3$$

with an r^2 of .97. When $X_2 = 1$, the relevant equation is

$$Y_i = 30.07 + .275X_3$$

and when $X_2 = 0$, the regression equation becomes

$$Y_i = 4.47 + .275X_3$$

These functions are plotted in Fig. 12-16. What would you expect to happen to r^2 if all the data were simply pooled together and the dummy variable omitted?

Table 12-5 Data for dummy variable problem

Observation	*Sales* Y	*Period* X_2	*Advertising budget* X_3
1	7	0	10
2	40	1	49
3	15	0	29
4	55	1	76
5	11	0	34
6	35	1	30
7	20	0	50
8	42	1	34
9	15	0	60
10	45	1	50
11	28	0	74
12	45	1	70
13	30	0	90
14	36	1	20
15	33	0	100
16	36	1	10

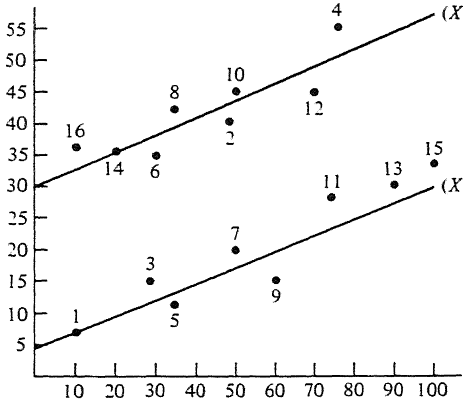


Fig. 12-16 Illustration of regression analysis with dummy variables.

LAGGED VARIABLES

In some situations it might be reasonable to hypothesize that changes in the dependent variable can best be explained *not* by current changes in the independent variable but by changes in the independent variable some one or two periods ago. Do you recall our example where sales of refrigerators was the dependent variable and housing starts the independent variable? Perhaps it might be more reasonable to suggest that the level of refrigerator sales in *this* period can best be explained by the level of housing starts *last period*. That is, a one-period lag is hypothesized. The rationale behind this is that it probably takes at least a period for housing starts to materialize into refrigerator orders. Therefore the paired observations that are used to determine the coefficients of the regression equation are sales in period t and housing starts in period $t - 1$. If, in fact, this is a better hypothesis than the one relating sales in period t with housing starts in period t, the coefficient of determination r^2 will be higher.

If the lagged relationship does prove to yield a higher r^2, a distinct advantage becomes evident when a prediction must be made. To predict next period's refrigerator sales, a lagged relationship of one period requires this period's housing starts, a figure that can be obtained quite easily.

In the absence of a lagged relationship, however, the prediction for next period's sales would require an estimate of next period's housing starts. Fortunately estimates for many economic variables do exist; however, if lagged relationships are effective, at least this source of error can be eliminated.

SUMMARY

The first stage in econometric model building is to identify a set of variables which can be used for predictive purposes. There is no rule to guide the model builder's selection of these variables, and in most cases he must rely on his intuition and insight. A considerable degree of "art" is involved at this stage.

Once these variables have been identified, the relevant data must be collected. The traditional source of these data has been government publications and trade journals. Consequently the data-gathering process has been very time-consuming. Several time-sharing companies, however, have changed all this. They offer data banks which have stored in them quarterly data on thousands of economic variables. To utilize these data the user need only specify the particular series of interest and the included years. In a multiple-regression run he might even specify several series from the data bank and several of his own. This new development has removed much of the drudgery from econometric analysis; several regressions can be run quite painlessly and in a short period of time.

In the second stage of the analysis a sound understanding of economic analysis is essential. Here the model builder must specify the form of the model, be it linear or nonlinear, perform the regression analysis, and interpret the results. He must watch for such problems as multicollinearity, autocorrelation, and heteroscedasticity.

Even if these problems are absent and the coefficient of determination is high, there still might be cause for rejection of the hypothesis. For it is absolutely necessary that the results make sense. If the signs in the regression equation are in the opposite direction from that which was expected, the result may be meaningless. Consider the hypothesis that refrigerator sales is a function of disposable income and housing starts. If in the regression equation the coefficient of housing starts is negative, the relationship which has been computed makes little sense. To propose that refrigerator sales decrease as housing starts increase is illogical. In conclusion, then, the regression relationship must not only pass statistical tests, but it must also make sense. Then again, should not all management science models pass this test?

QUESTIONS

Q12-1. What is the underlying difference between exponential smoothing and econometric methods.

Q12-2. What is the purpose of the disturbance term in the population regression relationship?

Q12-3. If in the case of simple linear regression all of the observations fell on a straight line, how would the sum-of-squares total and the sum-of-squares regression be related?

Q12-4. What is the consequence of fitting a linear equation to a nonlinear relationship?

Q12-5. Why should extra caution be exercised when extrapolating from a parabolic relationship?

Q12-6. What is heteroscedasticity, and how does one test for its presence?

Q12-7. What is autocorrelation, how does one test for its presence, and how is the problem resolved?

Q12-8. What problem exists when a high degree of correlation is present between independent variables in a multivariate relationship?

Q12-9. When is it appropriate to use dummy variables?

Q12-10. A manufacturer of fine men's clothes is interested in identifying the underlying causes of a salesman's success. Suppose that you accepted the project. What causative factors would you consider for inclusion in your model, and how would you collect the data?

Q12-11. Outline a demand theory for life insurance. What factors might affect the amount of insurance that an individual would purchase? Can these data be collected?

Q12-12. Identify those variables which a major TV manufacturer could incorporate into a regression model to forecast the demand for color TV sets.

Q12-13. In scheduling court cases it is important to estimate the length of time necessary to hear the case. What factors are likely to influence this time?

Q12-14. A few years ago the city of Boston was interested in playing host to the Bicentennial Celebration to be held in 1976. They commissioned a major consulting firm to generate a demand forecast based on the assumption that Boston would be the selected site.

The consulting firm decided to focus on individual states and to predict the attendance from each one of them. For historical data they obtained figures which had been collected for world fairs dating back to 1938.

If you were building the model, what causative factors would you include?

Q12-15. Specify a model which could be used to predict the value of a used car.

Q12-16. Identify the independent variables which you feel influence the level of crime. Could you collect the necessary data? Might your results be useful?

Q12-17. Identify the key factors which influence the demand for skiing.

PROBLEMS

P12-1. Compute the regression equation and coefficient of determination for the following data.

Y	X
2	10
10	20
7	32
15	38
17	50

P12-2. Compute the regression equation and coefficient of determination for the following data.

Y	X
1	15
3	20
2	18
5	30
10	45
8	40

P12-3. Compute the multiple-regression equation and coefficient of multiple determination for the following data.

Y	X_2	X_3
4	3	9
6	5	6
8	6	9
7	6	11
4	2	5
8	6	9
10	7	10
5	5	6

P12-4. To forecast aggregate demand for next year, the planning department of a large manufacturing firm has decided to employ econometric methods. The sales of one of their products, dishwashers, is felt to be sensitive to price and to the size of the advertising budget.

Given the following data, determine the regression relationship and the coefficient of determination.

Sales	*Price*	*Advertising*
17(000)	$18(0)	$28(000)
15	20	30
14	16	27
18	24	32
15	18	27
22	29	34
15	21	30
18	22	33
17	23	35
17	20	29

Generate a sales forecast for next year if the price they expect to charge is $200 and the advertising budget will be $30,000.

P12-5. Fit a second-degree parabola to the following data and compute the coefficient of determination.

Y	X
5	0
8	1
14	2
25	3
36	4

P12-6. Fit a straight line to the data given in Prob. 12-5. Compare the values of r^2.

BIBLIOGRAPHY

Aigner, Dennis: *Basic Econometrics,* Prentice-Hall, Inc., Englewood Cliffs, N.J., 1971.

Chambers, John C., Satinder K. Mullick, and Donald D. Smith: "How to Choose the Right Forecasting Technique," *Harvard Business Review,* July–August 1971, pp. 45–74.

Johnston, J.: *Econometric Methods,* McGraw-Hill Book Company, New York, 1963.

———: *Statistical Cost Analysis,* McGraw-Hill Book Company, New York, 1960.

Klein, Laurence R.: *An Introduction to Econometrics,* Prentice-Hall, Inc., Englewood Cliffs, N.J., 1962.

Parket, George G. C., and Edilberto L. Segura: "How to Get a Better Forecast," *Harvard Business Review,* March–April 1971, pp. 99–109.

Application Studies

Comanor, William S., and Bridger M. Mitchell: "Cable Television and the Impact of Regulation," *The Bell Journal of Economics and Management Science,* vol. 2, no. 1, Spring 1971, pp. 154–212.

Fisher, Franklin M., and Gerald Kraft: "The Effect of the Removal of the Fireman on Railroad Accidents, 1962–1967," *The Bell Journal of Economics and Management Science,* vol. 2, no. 2, Autumn 1971, pp. 470–494.

13

Aggregate Planning

INTRODUCTION

If demand for the firm's product or service were constant, the planning phase would be trivial. With demand at the same level day after day, week after week, a stable work force and production schedule could be established. Demand, however, is seldom constant. The demand for refrigerators, automobiles, hospital rooms, electric power, telephone services, classrooms, and so on usually displays substantial variability. The problem, then, is to develop a strategy by which these fluctuations in demand can be absorbed.

PLANNING FOR THE LONG, INTERMEDIATE, AND SHORT RUN

The strategy which is adopted will depend upon the *planning horizon:* long-range, intermediate, or short-range. If *long-range* plans are being formulated, there are no constraints on the planning process. If needed, new plants can be constructed, new equipment ordered, a labor force trained, and a new management team selected. In the long run, then, all input factors are variable. For some industries the long run may be any period over 1 year; in others it may be as much as 5 years. For example, the long run for the power industry is 5 years. It takes this long before the decision to increase capacity is followed by the opening of a new power plant.

It is within the constraint of previous long-range plans that *intermediate-range* plans are made. This horizon is usually considered to be from 1 month to 12 months. One of its distinguishing characteristics is that the physical plant and equipment are considered to be fixed over the planning horizon. Therefore fluctuations in demand cannot be met by changes in physical capacity but only through a strategy which might include some combination of overtime, hiring and layoff, inventory, and subcontracting. Each of these alternatives has a cost, and depending on the situation, some combination will absorb the fluctuations in demand most economically.

Planning for this intermediate range has received considerable emphasis, and as a result some insightful models have been developed. This chapter will focus on these models.

Short-range plans are formulated within the constraints of intermediate-range plans. By *short range* we usually mean between the next day and the next month. Usually the detail is much greater than in the aggregate plan.

The problem in short-range planning is to meet the imminent demand or commitments as economically as possible. The flexibility to accomplish this, however, is severely limited. Only a few strategies can be employed. These include the use of overtime, undertime, shifts in man and machine assignments, and changes in the sequence of jobs processed through the system.

The task of sequencing focuses on this short-range problem, and we turn to this issue in a later chapter.

AGGREGATE PLANNING

The essence of the intermediate-range planning problem is to devise a strategy by which fluctuations in demand can be economically absorbed. Demand, however, is usually specified in aggregate quantities. For example, if a firm manufactured 400 kinds of paint, the aggregate demand would be measured by the total number of gallons sold without regard to the product mix that made up this figure.

The first step in this process is to prepare a forecast for the aggregate level of demand. This may be in gallons of paint, number of cars, number of hospital beds, number of TV sets, and so on. Then with the help of a model, a plan is developed which specifies the work force size, production level, overtime requirements, and amount of subcontracting. All of these figures are on the aggregate level and serve as a basis for allocating resources to the firm's products when short-term scheduling is undertaken.

PURE STRATEGIES

There are several ways in which fluctuations in demand can be absorbed. In fact, some of them have already been suggested. Let us take a closer look at these alternatives and their associated costs. The first pure strategy involves *varying the production rate* in conformity with fluctuations in demand while keeping the size of the work force constant. One way to accomplish this is by utilizing overtime and undertime to absorb these fluctuations. Another way is through shorter workweeks during periods of slack demand. The difference between undertime and shorter workweeks is that *undertime* implies either a number of idle hours per week or a portion of the work force idle. Employees, however, are receiving full pay. A shortened workweek implies a less-than-40-hour workweek with less than 40-hour pay.

Each of these alternatives has an associated cost. The marginal cost for overtime is usually the easiest to estimate. Undertime and shortened workweeks, on the other hand, have costs which are not quite so easy to determine, for associated with these strategies is usually some attrition of the work force. Measuring this in dollars is quite difficult.

The second pure strategy is to *vary the size of the work force* in conformity with fluctuations in demand. This would be accomplished by hiring and layoff in direct relation to fluctuations in demand. These alternates also have associated costs. Hiring costs include employment interviews and selection, new records, physical examinations, and training. In addition, once the employee is on the job, it may take a substantial length of time before his productivity is on a par with that of the others who perform the same function. Layoff costs include those associated with unemployment compensation, insurance, severance pay, and intangible effects on public relations.

The third pure strategy is to maintain both a constant work force and a constant production level while absorbing fluctuations through *inventory* accumulation and depletion. This means that during times of below-average demand, excess production is accumulated in inventory. When demand is above average, it is met through inventory depletion. This strategy does not come free either. The cost includes the associated insurance, taxes, handling, obsolescence, and capital costs for carrying this extra inventory.

In addition to absorbing fluctuations through inventory accumulation and depletion, stockouts might be utilized. In periods of above-average demand and after inventory is depleted, it might be economical for further demand to result in a stockout. A stockout can lead to either a lost sale or a backorder. In the case of a lost sale the customer goes elsewhere to satisfy his demand, whereas in the case of a backorder the customer is willing to wait until the item is once again in stock. However difficult it is to estimate the cost of a stockout, it represents a viable way to absorb fluctuations in demand and must be considered in formulating a planning strategy.

Another pure strategy is to *subcontract* peak periods. In this way, the firm can set a constant production pace and turn to subcontractors to provide the extra capacity when required. The cost of this alternative will be the amount by which the subcontracting cost is greater than the marginal manufacturing cost at the high level of production.

When an organization must offer quick service and it is unable to inventory its product, another pure strategy can be taken. Examples of this can be found in the power and telephone companies. There is no way in which electric power or telephone switching capacity can be stored for use in peak periods. The only way that effective service can be offered is to design for peak use and *vary the utilization of capacity* according to off-peak demands. In general, then, some service industries absorb fluctuations in demand by varying the utilization of capacity.

We then have a set of pure strategies which management can utilize in absorbing fluctuations in demand. These include:

1. Overtime and undertime
2. Hiring and layoff
3. Inventory
4. Backorders
5. Subcontracting
6. Utilization of capacity

In most cases it is some *combination* of these pure strategies rather than a single strategy that will result in the most economical plan. Much of this chapter will now be devoted to building models which take into consideration different sets of these strategies in developing an aggregate production plan. First, we turn to the simplest of these models, one which takes into consideration only overtime and undertime, inventory, and subcontracting as alternatives for absorbing demand fluctuations. Then we will look at a series of two more models which consider different combinations of these strategies. Next a model is presented in which the costs associated with these strategies are nonlinear. Following this, a heuristic production-planning model is presented, and finally some additional planning strategies are suggested.

ABSORBING DEMAND FLUCTUATIONS WITH OVERTIME, UNDERTIME, INVENTORY, AND SUBCONTRACTING

Consider the case of the Quality Furniture Company. By using econometric forecasting techniques, a forecast for aggregate demand has been made. The results are shown in Table 13-1. Management has decided to focus on three

Table 13-1 Forecast of aggregate demand

Period	*Demand (000 units)*	*Period*	*Demand (000 units)*
1	10	7	12
2	15	8	10
3	30	9	18
4	27	10	26
5	30	11	30
6	16	12	17

alternatives for absorbing fluctuations in demand. These include overtime and undertime, inventory, and subcontracting. In the past they have relied heavily on local woodworking shops to absorb some of the manufacturing load in periods of peak demand. They have found subcontracting to be a convenient and economical alternative. This is not true of stockouts. They consider the cost of a stockout as being very high and therefore do not consider this alternative as a viable means of absorbing fluctuations in demand.

Recently they have asked their management science department to model the production-planning process. The first step taken by this department was to identify the costs associated with each of the smoothing strategies. These and the other parameters necessary for the analysis are presented in Table 13-2. The maximum number of units that can be produced on regular time is 19, and the maximum number that can be produced on overtime is 4 units. Regular-time production costs are $30 per unit, while overtime production costs are $35 per unit. Subcontractors will manufacture as many units as desired at a cost of $37 per unit. Finally, the inventory carrying costs were determined to be $1 per unit per period.

To develop a production plan for the 12-month period they simply proceeded to assign demand to production periods in the most economical way. To accomplish this, however, it is necessary to keep track of the period in which the unit will be manufactured or subcontracted and the period in which it will be sold.

Consider the first period, with a demand of 10 units. The most economical way to meet this demand is to produce the 10 units on regular-time production in period 1. This is recorded in the following way:

Period	*Demand*	*Reg.-time production*	*Overtime production*	*Subcontract*
1	10	10(1)		

To keep track of the production period, the quantity to be produced is entered in the row associated with period 1. Then a parenthesis is added in which is specified the sales month. The above entry can therefore be interpreted as regular-time production in period 1 of 10 units for sale in period 1.

Table 13-2 Production costs and constraints

Maximum regular-time production/period	19 units
Maximum overtime production/period	4 units
Cost of regular-time production	$30/unit
Cost of overtime production	$35/unit
Cost of subcontracting	$37/unit
Inventory carrying cost/period	$1/unit

Table 13-3 Simple linear cost model

Period	*Demand*	*Regular-time production*	*Overtime production*	*Subcontract*
1	10	10(1), 7(3), 2(4)		
2	15	15(2), 4(3)		
3	30	19(3)	2(4), 2(5)	
4	27	19(4)	4(4)	
5	30	19(5)	4(5)	5(5)
6	16	16(6)		
7	12	12(7), 7(11)		
8	10	10(8), 6(10), 3(11)		
9	18	18(9), 1(10)		
10	26	19(10)		
11	30	19(11)	1(11)	
12	17	17(12)		

Proceeding to period 2, the most economical way to meet the demand for 15 units is to produce 15 on regular-time production in period 2. This entry is made in Table 13-3. Period 3 demand is 30 units, and for the first time demand cannot be met by utilizing only regular-time production capacity. Moving from lower-cost alternatives to higher ones, 19 of the 30 units are assigned to regular-time production in period 3. This leaves 11 units of unscheduled demand that could be met by

1. Use of overtime in period 3 at a cost of $35 per unit
2. Subcontracting in period 3 at a cost of $37 per unit
3. Using the remaining units of regular-time capacity in period 2 at a cost of $30 per unit and storing in inventory for one period at a cost of $1 per unit, for a total cost of $31 per unit
4. Using the remaining units of regular-time capacity in period 1 at a cost of $30 per unit and storing in inventory for two periods, for a total cost of $32 per unit

The most economical approach is to choose the lowest-cost alternative in such quantities until a constraint is reached. If after this there is still unscheduled demand, the next most attractive alternative is used until its capacity is reached, and so on. Therefore, to meet the 11 units of unscheduled demand, 4 units are assigned to regular-time production in period 2. Since this fully allocates the regular-time capacity in period 2, 7 units still remain unassigned. The best remaining alternative is to assign these 7 units to period 1 regular time at a cost of $30 + $2 = $32 per unit.

Demand in period 4 is 27 units. Again the least-cost assignment begins with 19 units scheduled for regular-time production. The regular-time constraint having been met, the next best alternative is to assign 2 units to period

Table 13-4 Production plan

Period	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*
Regular-time prod.	19	19	19	19	19	16	19	19	19	19	19	17
Overtime prod.			4	4	4							
Subcontract					5							

1 regular-time production; this fully allocates the available regular-time production of period 1. So far 21 units have been assigned, leaving 6 units yet to be allocated. Four units can be placed in period 4 overtime at a cost of \$35 per unit, and the last 2 units can be allocated to period 3 overtime at a cost of \$35 + \$1 = \$36 per unit.

Demand in period 5 is 30 units. Nineteen units are again assigned to regular-time production in that period. Then 4 units are assigned to period 5 overtime, with 7 units yet to be assigned. Next, 2 units are assigned to overtime in period 3 at a cost of \$35 + \$2 = \$37. Finally, 5 units must be subcontracted in period 5 at \$37 per unit, as this is the most economical alternative remaining.

The allocation of demand continues for the remaining periods in the same manner. The optimal production plan thus determined is summarized in Table 13-4. The cost for this plan is \$7,410.

This model is relatively simple, since it considered only three alternatives for absorbing fluctuations in demand. These included regular and overtime production, inventory, and subcontracting. In addition the cost functions were all assumed to be linear. In the following models we will increase the number of alternatives which can be used for absorbing these demand fluctuations and also consider the consequence of nonlinear cost functions. Next we turn to a model which incorporates overtime and undertime, inventory, subcontracting, and backorders.

ABSORBING DEMAND FLUCTUATIONS WITH OVERTIME, UNDERTIME, INVENTORY, SUBCONTRACTING, AND BACKORDERS

Transportation models have been used to solve large-scale production-planning problems. The model that will be developed here will expand on the model presented in the last section in that backorders as well as beginning and ending inventories will be allowed. As in the previous model, a requirement of the transportation format is that we keep track of the period in which the unit is manufactured or subcontracted and the period for which it is demanded.

In Fig. 13-1 a transportation table for a three-period problem is illustrated. The top row represents initial inventory levels and the remaining rows represent regular-time, overtime, and subcontracting alternatives for each of the periods in the planning horizon. The level of initial inventory I_0 is entered

Sales period

			3	Final inventory	Slack	Capacity
Initial inventory	0	c	$2c$	$3c$	0	
Regular 1	r	$r+c$	$r+2c$	$r+3c$	0	R_1
Overtime 1	v	$v+c$	$v+2c$	$v+3c$	0	O_1
Subcontract 1	t	$t+c$	$t+2c$	$t+3c$	0	S_1
Regular 2	$r+b$	r	$r+c$	$r+2c$	0	
Overtime 2	$v+b$	v	$v+c$	$v+2c$	0	O_2
Subcontract 2	$t+b$	t	$t+c$	$t+2c$	0	S_2
Regular 3	$r+2b$	$r+b$	r	$r+c$	0	
Overtime 3	$v+2b$	$v+b$	v	$v+c$	0	
Subcontract 3	$t+2b$	$t+b$	t	$t+c$	0	
Demand	D_1	D_2	D_3	I_f	B	

Fig. 13-1 Transportation matrix.

I_0 = initial inventory; R_i = regular-time capacity in period i; O_i = overtime capacity in period i; S_i = quantity that can be subcontracted in period i; D_i = demand in period i; r = regular-time production cost per unit of output; v = overtime production cost per unit of output; t = subcontracting cost per unit of output; b = backorder cost per unit per period of backorder; c = inventory carrying charge per unit per period; I_f = final inventory; B = difference between total capacity and total demand when capacity is greater than demand.

in the last column of the first row. Below this are added the capacities of regular-time production, overtime production, and subcontracting for each period. Although in some cases regular-time production capacity may be the same for each period in the planning horizon, it is quite possible for it to vary from one period to the next. Therefore R_1 may not equal R_2, and R_2 may not equal R_3. This also holds for overtime and subcontracting capacity in each period.

The first three columns represent periods in the planning horizon: in this case, three. At the bottom of these columns is found the forecasted demand of each period. Another column is added if it is necessary to enter the next planning period (4) with some inventory. It is labeled the "final inventory" column and I_f, which is located at the bottom of this column, represents the desired level of this inventory.

The total capacity of the planning horizon is defined as the sum of the initial inventory, the regular-time, overtime, and subcontracting capacities for each period.

$$\text{Total capacity } I_0 + \sum_{i=1}^{3} R_i + \sum_{i=1}^{3} O_i + \sum_{i=1}^{3} S_i$$

The total demand, on the other hand, is equal to the sum of the period demands plus the desired level of final inventory.

$$\text{Total demand} = \sum_{i=1}^{3} D_i + I_f$$

To solve a transportation problem it is essential that total capacity equal total demand. When capacity is greater than demand, a slack demand column must be entered whose demand B is equal to the difference between total capacity and total demand.

$$B = \text{total capacity} - \text{total demand}$$

After B is introduced, the last row and last column should sum to the same grand total.

Now we turn to the cost entries in the body of the table. In the first row, the initial stock of inventory can be assigned to demand in any of the first three periods, final inventory, or slack. If a unit is assigned to the first period's sales, the marginal cost of this strategy is zero, since the unit has already been manufactured. If it is carried for this period and sold in the second, the marginal cost is c, the cost to carry one unit in inventory for one period. If it is carried into the third period, the cost is $2c$; and if it is carried into final inventory, the cost is $3c$, since the unit will have been carried for three full periods. Finally an entry in the slack column would imply that the unit of initial inventory is not used at all, and the marginal cost for this strategy would be zero.

If a unit of demand in period 1 is assigned to regular-time production in period 1, the production cost is r. If, on the other hand, a unit of demand

Table 13-5 Three-period production-planning problem

I_0 = 40 units	S_1 = 15 units	v = \$60/unit
R_1 = 50 units	S_2 = 25 units	t = \$67/unit
R_2 = 65 units	S_3 = 30 units	b = \$2/unit/period
R_3 = 80 units	D_1 = 150 units	c = \$10/unit/period
O_1 = 20 units	D_2 = 60 units	I_f = 70 units
O_2 = 25 units	D_3 = 80 units	
O_3 = 30 units	r = \$40/unit	

in period 2 is assigned to regular-time production in period 1, the cost is $r + c$, since the unit must be carried in inventory for one period at a cost of c. If a unit of demand from period 3 is assigned to regular-time production in period 1, the cost is $r + 2c$; and if a unit of final inventory is assigned to this production period, the cost is $r + 3c$. Again we have a cost of zero if a unit of slack demand is assigned to regular-time production in period 1. Units assigned to that cell imply unused capacity for regular-time production in period 1. The next two rows in the matrix can be interpreted in a similar fashion, except that v represents the per-unit cost of overtime and t the per-unit cost of subcontracting.

Next we turn to production in period 2. If demand in period 1 is assigned to regular-time production in period 2, it should be clear that a backorder is implied. If backorder costs are b per unit per period, the cost for this period is $r + b$. When demand in period 2 is assigned to regular-time production in period 2, the cost becomes r. There should be little difficulty interpreting the rest of the costs associated with period 2.

In period 3, when demand in period 1 is assigned to regular-time production in period 3, backorder costs for two periods are incurred. Therefore the cost for this cell becomes $r + 2b$. The rest of the costs for period 3 should now be in a familiar sequence.

Data are given in Table 13-5 for a three-period production-planning problem. The matrix is illustrated in Fig. 13-2 and the solution is shown in Fig. 13-3.

ABSORBING DEMAND FLUCTUATIONS WITH OVERTIME, INVENTORY, HIRING, AND LAYOFF

In our next production-planning model two new strategies will be introduced. These include hiring and layoff. Through them we will be able to impose a cost on increasing or decreasing the production level in any period.

Our two previous models have allowed production levels to increase or decrease within the upper bound of regular-time production capacity, but no costs were incurred when these *changes* were made. For example in the solution to the problem presented in Table 13-3, regular-time production is adjusted in periods 6, and 12. How was this change accomplished? If hiring and layoffs

	Sales period					
			3	Final inventory	Slack	Capacity
Initial inventory	0	10	20	30	0	40
Regular 1	40	50	60	70	0	50
Overtime 1	60	70	80	90	0	20
Subcontract 1	67	77	87	97	0	15
Regular 2	42	40	50	60	0	65
Overtime 2	62	60	70	80	0	25
Subcontract 2	69	67	77	87	0	25
Regular 3	44	42	40	50	0	80
Overtime 3	64	62	60	70	0	30
Subcontract 3	71	69	67	77	0	30
Demand	100	60	130	70	20	380

Fig. 13-2 Production-planning problem in transportation format.

were employed, certainly some costs were incurred. It is these costs that will be introduced in our new model.

In addition to hiring and layoff the other variables included will be overtime and inventory. Our objective will be to minimize the sum of regular-time production costs, hiring, layoff, overtime, and inventory carrying costs. The solution will be constrained, however, by regular-time and overtime production capacity, the level of hiring and layoff, and the necessity to meet forecasted demand.

	Sales period					
				Final inventory	Slack	Capacity
Initial inventory	0 40	10	20	30	0	40
Regular 1	40 50	50	60	70	0	50
Overtime 1	60 20	70	80	90	0	20
Subcontract 1	67 15	77	87	97	0	15
Regular 2	42 25	40 35	50	60 5	0	65
Overtime 2	62	60 25	70	80	0	25
Subcontract 2	69	67	77	87 5	0 20	25
Regular 3	44	42	40 80	50	0	80
Overtime 3	64	62	60	70 30	0	30
Subcontract 3	71	69	67	77 30	0	30
Demand	150	60	80	70	20	380

Fig. 13-3 Solution matrix.

In the interest of keeping the model to a manageable size, a three-period model will be developed where:

D_i = predicted demand in period i
P_i = scheduled regular-time production in period i
M_i = *maximum* regular-time production that can be scheduled in period i. This is based on an estimate made at the beginning of the planning horizon and usually represents the limitation imposed by physical plant capacity.

T_i = overtime scheduled in period i
Y_i = *maximum* overtime production that can be scheduled in period i. This is based on an estimate made at the beginning of the planning horizon.
v = overtime cost per unit
c = inventory carrying charge per unit per period
h_i = cost of increasing the production level by one unit of output (hiring)
f_i = cost of reducing the production level by one unit of output (layoff)
A_i = increase in production level in period i
R_i = reduction in production level in period i
r = regular-time production cost per unit

First we will develop the relevant constraints and then the objective function.

REGULAR-TIME PRODUCTION CONSTRAINTS

The first set of constraints specifies the limit to the regular-time production that can be scheduled in each period. In the first period this can be expressed as

$$P_1 \leq M_1$$

where the scheduled regular-time production P_1 must be less than or equal to the regular-time production capacity. The regular-time production capacity is required as input data to the model and must therefore be estimated.

In the second period the regular-time production constraint can be written in the following way:

$$P_2 \leq M_2$$

Again, the estimate for M_2 is made at the beginning of the planning horizon. The constraint then ensures that the scheduled regular-time production in period 2 will be less than or equal to the previously estimated regular-time capacity. Following directly from this pattern, the constraint for the third period can be written as

$$P_3 \leq M_3$$

Next we turn to constraints on overtime production.

OVERTIME PRODUCTION CONSTRAINTS

Once again capacity estimates must be made beforehand. This time they include the maximum overtime production that can be scheduled in each period. The overtime constraint for the first period can then be expressed as

$$T_1 \leq Y_1$$

where T_1 = scheduled overtime in period 1
Y_1 = previously estimated overtime production capacity in period 1

The constraint for period 2 can be expressed as

$$T_2 \leq Y_2$$

where Y_2 = previously estimated maximum overtime capacity for period 2

Following the same approach, the constraint for period 3 can be written as

$$T_3 \le Y_3$$

Next we turn to the inventory constraints.

INVENTORY CONSTRAINTS

If we assume that the initial level of inventory is zero, the inventory constraint for the first period can be written in the following way:

$$P_1 + T_1 \ge D_1$$

That is, regular-time production in period 1 plus overtime production in period 1 must at least meet forecasted demand for period 1. In period 2 we have the following:

$$P_1 + T_1 - D_1 + P_2 + T_2 \ge D_2$$

or

$$\sum_{i=1}^{2} P_i + \sum_{i=1}^{2} T_i \ge \sum_{i=1}^{2} D_i$$

This specifies that the total production of period 1 less demand for that period plus the total production of period 2 must be at least as great as demand in period 2. The inventory constraint for the first period ensured that production must be *at least* as great as demand. Combined with the second constraint it allows inventory to be carried over into period 2 so that demand in period 2 can be met from ending inventory in period 1, $P_1 + T_1 - D_1$, and production in period 2, $P_2 + T_2$. Continuing with the inventory constraint for period 3, we have

$$\sum^{3} P_i + \sum_{i=1}^{3} T_i \ge \sum_{i=1}^{3} D_i$$

HIRING CONSTRAINTS

In the next two sets of constraints the extent of the production level changes is identified. First we look at production level increases (hiring). In the first period this constraint can be formulated in the following way:

$$A_1 \ge P_1 - P_0$$

That is, the increase in the production level A_1 must be greater than or equal to the difference between scheduled production in period 1 and that in the preceding period. The value of P_0 must be supplied as data to the model. Since the period which precedes the planning model is often the current period, the estimation of P_0 is generally no problem.

In the second period we have

$$A_2 \ge P_2 - P_1$$

and finally in the third period

$$A_3 \geq P_3 - P_2$$

LAYOFF CONSTRAINTS

Next the extent of any reductions in the production levels is identified. In the first period we have

$$R_1 \geq P_0 - P_1$$

where the reduction must be greater than or equal to the decrease in the production level between successive periods. For periods 2 and 3 we have

$$R_2 \geq P_1 - P_2$$

$$R_3 \geq P_2 - P_3$$

OBJECTIVE FUNCTION

Now we turn to the objective function, where the objective will be to minimize the sum of the costs incurred over the three-period planning horizon. The costs incurred in the first period include the following components:

$$rP_1 + hA_1 + fR_1 + vT_1 + c(P_1 + T_1 - D_1)$$

Now we turn to the objective function, where the objective will be to minimize the sum of the costs incurred over the three-period planning horizon. In the second period we have

$$rP_2 + hA_2 + fR_2 + vT_2 + c[(P_1 + T_1 - D_1) + (P_2 + T_2 - D_2)]$$

where the sum in the bracket represents the inventory carried into period 3. Costs in the third period are

$$rP_3 + hA_3 + fR_3 + vT_3 + c[(P_1 + T_1 - D_1) + (P_2 + T_2 - D_2) \\ + (P_3 + T_3 - D_3)]$$

Now we can write the objective function as the minimization of the sum of these costs.

$$\text{Minimize cost} = r \sum_{i=1}^{3} P_i + h \sum_{i=1}^{3} A_i + f \sum_{i=1}^{3} R_i + v \sum_{i=1}^{3} T_i \\ + c \left[\sum_{i=1}^{3} \sum_{j=1}^{i} (P_j + T_j - D_j) \right]$$

Summarizing the model in general notation we have

$$\text{Minimize cost} = r \sum_{i=1}^{k} P_i + h \sum_{i=1}^{k} A_i + f \sum_{i=1}^{k} R_i + v \sum_{i=1}^{k} T_i \\ + c \left[\sum_{i=1}^{k} \sum_{j=1}^{i} (P_j + T_j - D_j) \right]$$

$$P_i \leq M_i \qquad i = 1, 2, 3, \ldots, k$$

$$T_i \leq Y_i \qquad i = 1, 2, 3, \ldots, k$$

$$\sum_{j=1}^{i} P_j + \sum_{j=1}^{i} T_j \geq \sum_{j=1}^{i} D_j \qquad i = 1, 2, 3, \ldots, k$$

$$A_i \geq P_i - P_{i-1} \qquad i = 1, 2, 3, \ldots, k$$

$$R_i \geq P_{i-1} - P_i \qquad i = 1, 2, 3, \ldots, k$$

All $A_i, R_i, T_i, P_i \geq 0$.

Example Consider the following example:

$D_1 = 100$ units

$D_2 = 50$

$D_3 = 75$

$v = \$15/\text{unit}$

$c = \$5/\text{unit}$

$h = \$30/\text{unit}$

$M_1 = 90$ units

$M_2 = 60$

$M_3 = 60$

$f = \$10/\text{unit}$

$r = \$10/\text{unit}$

$P_0 = 75$

$Y_1 = 15$ units

$Y_2 = 10$

$Y_3 = 10$

The objective function and constraints can be written in the following way:

$$\text{Minimize cost} = 10 \sum_{i=1}^{3} P_i + 30 \sum_{i=1}^{3} A_i + 10 \sum_{i=1}^{3} R_i + 15 \sum_{i=1}^{3} T_i + 5 \left[\sum_{i=1}^{3} \sum_{j=1}^{i} (P_j + T_j - D_j) \right]$$

$$P_1 \leq 90$$

$$P_2 \leq 60$$

$$P_3 \leq 60$$

$$T_1 \leq 15$$

$$T_2 \leq 10$$

$$T_3 \leq 10$$

$$P_1 + T_1 \geq 100$$

$$P_1 + P_2 + T_1 + T_2 \geq 150$$

$$P_1 + P_2 + P_3 + T_1 + T_2 + T_3 \geq 225$$

$$A_1 \geq P_1 - 75$$

$$A_2 \geq P_2 - P_1$$

$$A_3 \geq P_3 - P_2$$

$$R_1 \geq 75 - P_1$$

$$R_2 \geq P_1 - P_2$$

$$R_3 \geq P_2 - P_3$$

All $A_i, R_i, T_i, P_i \geq 0$.

Using a computer with a standard linear programming code, the results were found to be:

$P_1 = 85$ $T_1 = 15$ $A_1 = 10$ $R_1 = 0$
$P_2 = 60$ $T_2 = 0$ $A_2 = 0$ $R_2 = 25$
$P_3 = 60$ $T_3 = 5$ $A_3 = 0$ $R_3 = 0$

A MORE ADVANCED LP MODEL

A more advanced linear programming model for production planning has been developed by Hanssmann and Hess.[1] In addition to the overtime, hiring and layoff, and inventory variables which we just included, they also include undertime and shortages as alternative means of absorbing demand fluctuations. Furthermore their model does not require that estimates be made of regular-time and overtime capacities. They therefore assume that there have been no commitments made ahead of time which affect the regular-time and overtime capacity in each future production period. Rather, their model sets these production levels directly.

ABSORBING DEMAND FLUCTUATIONS WHEN OVERTIME, UNDERTIME, HIRING, LAYOFF, INVENTORY, AND BACKORDER COSTS ARE NONLINEAR

A major assumption of the planning models developed thus far has been that the relevant costs were linear. That is, the cost of carrying the first unit in inventory is the same as the cost of carrying the one-thousandth unit in inventory. The cost of hiring the first employee is the same as the cost of hiring the one-hundred-first employee, and so on. It might be reasonable to accept these linear assumptions for some aggregate planning problems; however, in others it might not be appropriate to accept them but would be better to develop a nonlinear model instead.

The nonlinear model developed by Holt, Modigliani, and Simon[2] is the best-known. In their model, quadratic functions are used to approximate regular payroll, hiring and layoff, overtime and undertime, and inventory and backordering costs. To gain some insight into their approach we will develop their quadratic cost functions and their solution and illustrate an application which they documented in their classic paper.

REGULAR PAYROLL COST

Their model assumes that regular payroll commitments are made for one month at a time. Therefore regular payroll costs for any month t are

$$C_1 W_t$$

---

[1] F. Hanssmann and S. W. Hess, "A Linear Programming Approach to Production and Employment Scheduling," *Management Technology,* Shore: *Operations Management,* p. 603, 1, January 1960.

[2] C. C. Holt, F. Modigliani, and H. A. Simon, "A Linear Decision Rule for Production and Employment," *Management Science,* vol. 2, no. 2, October 1955.

where C_1 = monthly wage rate
W_t = number of employees in work force during month t

This regular payroll cost is illustrated in Fig. 13-4.

HIRING AND LAYOFF COSTS

Hiring and layoff costs vary not with the size of the work force but with changes in its size. An estimate of how these costs behave is shown as the solid line in Fig. 13-5. When hiring is undertaken, $W_t - W_{t-1} > 0$, and linear hiring and training costs are incurred. On the other hand, if layoffs are incurred, $W_t - W_{t-1} < 0$, and linear layoff costs are incurred. A suitable approximation to this function is the quadratic curve shown in Fig. 13-5 as a dotted line. In fact the interpretation that goes along with this quadratic function seems very reasonable. That is, as the quantity of workers laid off or hired increases, costs rise at an increasing rate. For example, the efficiency of hiring, perhaps measured by the quality of new-hires, decreases as the quantity hired at one time increases. As for layoffs, reorganization costs increase more than proportionately for large layoffs as compared with small ones. Even if these arguments do not hold, the authors claim that the quadratic is a good approximation over the usual ranges of interest. Mathematically the quadratic function can be approximated as $C_2(W_t - W_{t-1})^2$, where C_2 is chosen to ensure a good statistical fit.

OVERTIME COSTS

Overtime costs are incurred when production levels exceed the regular-time productive capacity of the work force. These costs depend on two variables: the

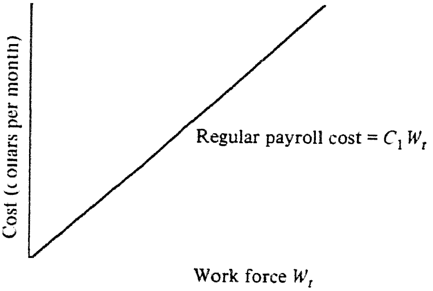


Fig. 13-4 Regular payroll costs. (*From C. C. Holt et al., Planning Production, Inventories, and Work Force, Prentice-Hall, Inc., Englewood Cliffs, N.J., 1960, p. 8.*)

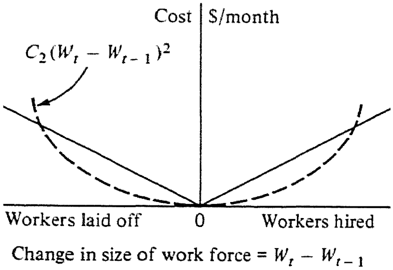


Fig. 13-5 Hiring and layoff costs. (*From C. C. Holt et al., Planning Production, Inventories, and Work Force, Prentice-Hall, Inc., Englewood Cliffs, N.J., 1960, p. 8.*)

size of the work force at a given time and the aggregate production rate. For a given production rate, the larger the work force level, the less the overtime required.

The quadratic function illustrated in Fig. 13-6 represents an approximation of these costs for a particular work force level W_t. For larger work force levels there are curves which lie to the right and for smaller levels curves which lie to the left. From this figure we can see that, as production increases, overtime costs slowly increase and then proceed to increase faster and faster. This can be attributed to the fact that as capacity is reached, a few bottlenecks occur and only a few employees must work overtime. As production increases further, more employees are required, until the whole work force is on overtime. This function can be approximated in the following way:

$$C_3(P - C_4W_t)^2 + C_5P - C_6W_t$$

where C_4 represents the productivity of the work force and both C_5 and C_6 are added to improve approximation.

INVENTORY AND BACKORDER COSTS

The final set of alternatives that are considered for absorbing demand fluctuations includes inventory and backorder costs. The ideal aggregate level of inventory is determined by summing the optimal level for each product carried in inventory. The methods for determining these optimal levels will be covered in the next chapter. For purposes of this chapter it is enough to know that the ideal aggregate level can be calculated and that actual levels which deviate from this ideal incur added costs. Inventory levels which are higher incur those costs that are associated with carrying extra inventory, whereas levels that are

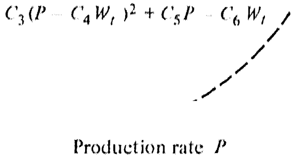


Fig. 13-6 Overtime costs. (*From C. C. Holt et al., Planning Production, Inventories, and Work Force, Prentice-Hall, Inc., Englewood Cliffs, N.J., 1960, p. 10.*)

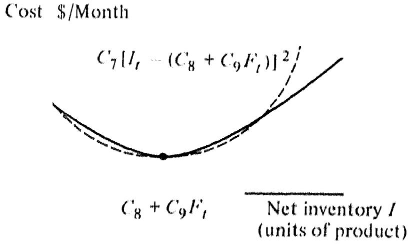


Fig. 13-7 Inventory and backorder costs. (*From C. C. Holt et al., Planning Production, Inventories, and Work Force, Prentice-Hall, Inc., Englewood Cliffs, N.J., 1960, p. 13.*)

lower incur extra shortage costs, since the likelihood of running out of stock increases.

The optimal inventory level can be approximated as $C_8 + C_9F_t$, where F_t is the order rate (demand) in period t. These inventory-related costs are shown as the solid line in Fig. 13-7 and can be approximated by a quadratic shown as the dotted line. The form of the quadratic is

$$C_7[I_t - (C_8 + C_9F_t)]^2$$

where C_7, C_8, C_9 = costs statistically chosen to ensure good fit to data
I_t = inventory level in period t

TOTAL-COST FUNCTION

Now the individual cost components can be combined into a total-cost function. The objective will be to schedule production in such a way as to minimize production and employment costs. Since each month's decision has repercussions which extend into the future, several production periods must be considered. If we consider N periods, the model can be expressed as

$$\text{Minimize } C_N = \sum_{t=1}^{N} C_t$$

$$\begin{aligned} \text{where } C_t = {} & [(C_1W_t) && \text{Regular payroll costs} \\ & + C_2(W_t - W_{t-1})^2 && \text{Hiring and layoff costs} \\ & + C_3(P_t - C_4W_t)^2 + C_5P_t - C_6W_t && \text{Overtime costs} \\ & + C_7(I_t - C_8 - C_9F_t)^2] && \text{Inventory-connected costs} \end{aligned}$$

subject to the constraints

$$I_{t-1} + P_t - F_t = I_t \qquad t = 1, 2, \ldots, N$$

These constraints establish the relationship between inventory at the beginning of the month, production during the month, sales during the month, and ending inventory. Furthermore, they must hold for each of the N months.

THE MODEL APPLIED

This model was applied to data collected from a paint factory for the years 1949 to 1954. The estimates for the cost coefficients were statistically determined from accounting data, together with the subjective estimates of such intangible costs as delayed shipments to customers. The result was the following:

$$\begin{aligned} \text{Minimize } C_N = \sum_{t=1}^{N} \{ & (340W_t) + [643(W_t - W_{t-1})^2] \\ & + [.20(P_t - 5.67W_t)^2 + 51.2P_t - 281W_t] \\ & + [.0825(I_t - 320)^2]\} \end{aligned}$$

subject to

$$I_{t-1} + P_t - F_t = I_t \qquad t = 1, 2, \ldots, N$$

After these parameters were estimated, the next step was to obtain a solution. This was achieved by differentiating with respect to each decision variable. The result of this procedure is shown in Fig. 13-8. This represents the optimal decision rule for production and work force levels in period t.

The first rule that is applied at the beginning of the month is the production rule. It incorporates a weighted average of forecasts extending some 11 months into the future. The weight is heaviest on the most recent months and diminishes rapidly until the impact of forecasts beyond month 11 is so small that they can be ignored. This pattern of weights implies that orders well into the future have little impact on current production levels. In fact the next two months account for most of the influence on the level of current production.

The second term, $.993W_{t-1}$, represents the impact that employment levels in the previous month have on production levels in the current month. The greater the previous level of the work force, the higher will be current production. This reflects the costs of hiring and layoff together with undertime and overtime as costly means of absorbing fluctuations.

The last two terms reflect the effect that the previous month's inventory

$$P_t = \left\{ \begin{array}{l} +.463F_t \\ +.234F_{t+1} \\ +.111F_{t+2} \\ +.046F_{t+3} \\ +.013F_{t+4} \\ -.002F_{t+5} \\ -.008F_{t+6} \\ -.010F_{t+7} \\ -.009F_{t+8} \\ -.008F_{t+9} \\ -.007F_{t+10} \\ -.005F_{t+11} \end{array} \right. + .993W_{t-1} + (153 - .464I_{t-1})$$

$$W_t = .743W_{t-1} + (2.09 - .010I_{t-1}) + \left\{ \begin{array}{l} +.0101F_t \\ +.0088F_{t+1} \\ +.0071F_{t+2} \\ +.0054F_{t+3} \\ +.0042F_{t+4} \\ +.0031F_{t+5} \\ +.0023F_{t+6} \\ +.0016F_{t+7} \\ +.0012F_{t+8} \\ +.0009F_{t+9} \\ +.0006F_{t+10} \\ +.0005F_{t+11} \end{array} \right.$$

Fig. 13-8 Production and work force level rules.

has on the current production level. For a large ending inventory the sum of the two terms is negative and production will be decreased. In this way the rule responds to past forecasting errors.

The second decision rule is used to compute the size of the work force. The last set of terms reflects the impact that forecasts have on the current work force level. In this rule, however, the weights extend further into the future before becoming negligible in size. It can be concluded, therefore, that distant forecasts have more effect on the current work force level than they have on the production level.

The first term, $.743W_{t-1}$, reflects the impact that the work force level at the end of the preceding period has on the current work force decision.

The next two terms specify the effect that the previous level of inventory has on the current work force decision. A large inventory position leads to a decrease in work force level, whereas a small inventory leads to an increase. Inventory, however, has a smaller impact on work force levels than it does on production.

In addition to developing these rules for the paint factory, they were tested over a period of 5 years for which data were available. The rules were used to simulate the decisions that would have been made over this 5-year interval.

Before the test could proceed, however, a forecasting scheme had to be developed. A very simple technique known as a *moving average* was selected. An n-month moving average is simply an average of data for n months with one modification. With the passage of 1 month, the earliest observation is dropped, the most recent is added, and a new average is computed. Therefore the average will always include data for the most recent n months. In general, exponential smoothing and econometric analysis are much more sophisticated; but for purposes of testing this decision rule, a simple forecasting scheme was employed.

In addition to the test based on the moving average forecast, the decision rule was also tested with a perfect forecast which was based on orders actually received. This provided a good basis for comparison, since it set an upper limit to the performance which could have been achieved with the decision rule.

In all, a three-way comparison was performed. This included the actual performance of the factory, the performance of the decision rule when the moving average forecast was used, and the performance of the rule with a perfect forecast.

The production levels associated with these three alternatives are shown in Fig. 13-9. It can be seen that the actual month-to-month fluctuations in production were greater than those associated with either the perfect forecast or the moving average.

The cost performance for each of these alternatives was computed by using the quadratic cost functions. The results are summarized in Table 13-6. The decision rule which used the perfect forecast was only 10 percent lower than the rule with a moving average foreast and 39 percent lower than the actual performance of the paint company. This implies that even crude forecasts when

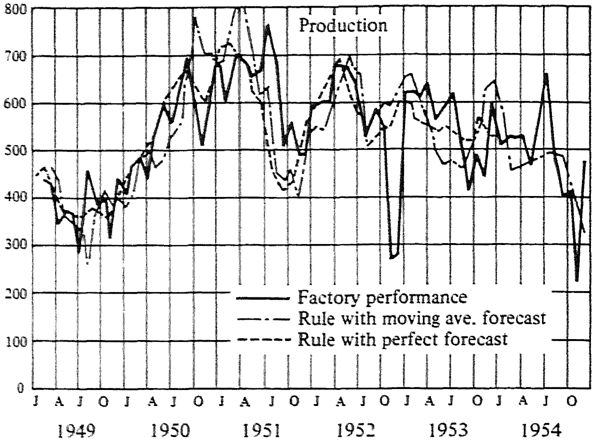


Fig. 13-9 5-year comparison. *(From C. C. Holt et al., Planning Production, Inventories, and Work Force, Prentice-Hall, Inc., Englewood Cliffs, N.J., 1960, p. 21.)*

combined with optimal decision rules can outperform actual company performance by a substantial measure. In fact, since the moving average performance was but 10 percent higher than that of the perfect forecast performance, it is questionable in this case whether a forecasting technique more sophisticated and more costly than a simple moving average is required. This, of course, would depend on the marginal cost of the better forecast and the marginal savings it generated over the moving average forecast.

Table 13-6 Results of five-year comparison

Costs (thousands of dollars)	*Company performance*	*Decision rule*	
		Moving avg. forecast	*Perfect forecast*
Regular payroll	$1,940	$1,834	$1,888
Overtime	196	296	167
Inventory	361	451	454
Backorders	1,566	616	400
Hiring and layoffs	22	25	20
Total cost	$4,085	$3,222	$2,929
	139%	110%	100%

Source: C. C. Holt, F. Modigliani, and H. A. Simon, "A Linear Decision Rule for Production and Employment," *Management Science*, vol. 2, no. 2, October 1955, p. 26.

A HEURISTIC MODEL FOR PRODUCTION PLANNING

Now that we have progressed from simple linear models to more complex linear models and finally to quadratic models, let us move out of the realm of optimization models and investigate a heuristic approach to production planning.

Bowman has proposed a method whereby the form of the decision rule is determined by rigorous analysis; however, the coefficients of the model are determined by regression analysis.[3] For example, consider the following very intuitive production rule:

$$P_t = K + aW_{t-1} - bI_{t-1} + cF_{t+1} \tag{13-1}$$

It implies that the production level in period t should depend upon the work force level in the previous period W_{t-1}, the inventory in the previous period I_{t-1}, and the forecast of demand in the following period F_{t+1}. Now, if management has been sensitive to these criteria in setting their production levels in the past, it seems reasonable to make use of their insight. Often, however, their shortcoming is such that their behavior tends to be quite variable but not off center. Bowman's contribution was to overcome this problem of variability by using regression analysis to determine the coefficients of the model. When the model is used, a more constant decision will be reached each period.

Gordon[4] collected data from a brewery on the variables presented in Eq. 13-1. For each period data were gathered on production level, work force level in the previous period, inventory level in the previous period, and the forecast for the next period's sales. A regression was run with the following outcome:

$$P_t = 6.98 + 1.66W_{t-1} - .12I_{t-1} + .44F_{t+1}$$

In addition to the production rule a work force rule was also derived:

$$W_t = 4.20 + .63W_{t-1} + .17F_{t+1}$$

These heuristic rules were tested and compared with the actual performance of the company. They showed an improvement.

Perhaps the biggest appeal of this approach is that it is patterned after the actual decision process. For some it is reassuring to know that the model is based on behavior patterns generally acceptable to the firm. It is necessary, however, to have the relevant historical data before the coefficients can be estimated. This may be impossible for new product lines or for situations in which the historical data are not available. In addition, the continued use of the regression model implies that no underlying change in the importance and impact of the independent variables is expected to occur. That is, relationships that held in the past will also hold in the future.

---

[3] E. H. Bowman, "Consistency and Optimality in Managerial Decision Making," *Management Science*, vol. 9, 1963, pp. 310–321.

[4] J. R. M. Gordon, "A Multi-model Analysis of an Aggregate Scheduling Decision," unpublished Ph.D. dissertation, Sloan School of Management, M.I.T., 1966. Report in Elwood S. Buffa, *Production Inventory Systems*, Richard D. Irwin, Inc., Homewood, Ill., 1968, p. 174.

SOME OTHER CONSIDERATIONS IN PRODUCTION PLANNING

Compared with the full range of possibilities for smoothing production, only part of the story has been told. Those strategies which we have neglected, however, are frequently considered as outside the domain of the operational focus and inside that of marketing, finance, and top administrative levels. It is necessary that we consider them here, as they are but other means for absorbing demand fluctuations and could well represent more economical strategies than some already presented.

PRICING AND PROMOTION

First we will consider a situation where demand can be influenced by pricing and promotional strategies.[5] In some cases the quantity of a product demanded is independent of the price charged and the level of promotion. This is referred to as *inelastic demand.* There are other products, however, whose quantity demanded is affected by these strategies. If the price is lowered, the quantity sold increases; and conversely, if the price is raised, the quantity sold decreases.

Promotional strategy may also account for changes in demand. As the promotional activity increases, the demand for the product may also increase.

When pricing and promotion affect quantity demanded, these strategies can be used to shift the peak demand to off-peak periods. Perhaps the classic examples of this are found in the public utilities. The demand for telephone communication reaches a peak during business hours and decreases rapidly into the evening hours and on weekends. To accommodate this peak demand, telephone switching equipment is required which goes unused during the off-peak hours. Rather than absorb all of this peak demand by varying the utilization of capacity, the telephone company has adopted a policy of differential pricing. Calls are most expensive during business hours, less expensive in the early evening hours and on weekends, and even less expensive between midnight and 7:00 A.M. The evidence shows that this pricing strategy has been quite effective in leveling peak demand.

Other examples of pricing strategies directed toward shifting demand can be found in the airline industry. During the peak summer weekends, overseas airline passengers are taxed an additional $50. Excursion fares which allow travel only during off-peak days are yet another example.

The materialistic capital of the world, Miami Beach, has used promotional strategies to increase demand during the off season. Many tourists now visit the area year-round. Of course, the peak is still from December to March.

ADAPTATIONS IN PRODUCT, PROCESS, LOCATION, OR POLICY

Consider a seasonal business in which it is impossible to avoid hiring and layoffs. To lower the costs associated with these strategies it may be quite effective to locate in a rural area where, for example, farmers might work in the factory during the winter months and attend to their farming during the summer. This strategy suits both employer and employee quite well, and as a result the imponderable costs usually associated with this approach are avoidable.

---

[5] Jay R. Galbraith, "Solving Production Smoothing Problems," *Management Science,* vol. 15, no. 12, August 1969, p. B-665.

The telephone company has a policy of hiring part-time women workers during peak hours of the day. The women do not object, since those who accept the job need the rest of the day for taking care of their families, shopping, etc.

The airlines have changed the design of their process to accommodate daily fluctuations. They have introduced planes which fly passengers during the day and, with minor changes, fly freight at night. No longer is it necessary to keep costly equipment idle at night just because people prefer flying during the day.

Finally, with the cooperation of those responsible for product development and design, new products could be added which have a contracyclical demand pattern. The classic example is the ice skate manufacturer who introduces a line of baseball gloves. The ideal, of course, is to add a product whose peak demand occurs during the valleys of the current product line.

Next we turn to two strategies which focus on the coordinated effort of a group of organizations to solve production smoothing problems. These include bargaining and coalition.

BARGAINING

On many occasions it is possible for the customer and supplier to coordinate their operations through the bargaining process. The supplier, for example, might offer price incentives or preferential treatment if the customer would accept his order in increased lots, split lots, delayed delivery, or faster delivery. This would give the supplier enough flexibility to shift some of the peak demand into the valleys and thereby increase the efficiency of his process.

COALITION

Most large electric power companies produce a good share of their power requirements. To ensure their ability to meet peak demands, however, they are tied into regional power grids. Such a grid unites several power companies and thereby provides a source for extra power. Therefore through the forming of this coalition their resources can be pooled, and the total investment required is reduced.

SUMMARY

The focus in this chapter has been twofold: first, the development of alternatives to absorb demand fluctuations; second, the use of models, where possible, to take into account the interactions between a set of relevant alternatives.

The alternatives from which the decision maker can choose include the following:

1. Overtime and undertime
2. Hiring and layoff

3. Inventory
4. Backorders
5. Subcontracting
6. Utilization of capacity
7. Pricing and promotion
8. Adaptations in product, process, location, or policy
9. Bargaining
10. Coalition

Each has its own cost. In fact the most difficult aspect of the production-planning process may be the estimation of these costs. It is quite unlikely that they can be read directly from accounting records.

The structure of this chapter was such that it led you from the simplest linear model to more complex linear models, to a nonlinear model, to a heuristic approach, and finally to a set of three other strategies which have not yet been included in the modeling process.

Once again the problem is to identify those factors in the decision environment which are relevant for purposes of the decision at hand. It may be that some or all of these factors may be included in a model. To the extent that some of them cannot be incorporated in the model, the decision maker is required to employ personal judgment when the final decision is made.

QUESTIONS

Q13-1. What is the difference between long-range, intermediate-range, and short-range planning?

Q13-2. Under what conditions should backorders be incurred?

Q13-3. What is the major appeal of the heuristic production-planning model?

Q13-4. How can pricing and promotion be used to smooth demand for an urban rapid transit system?

Q13-5. Discuss the benefits and problems of generating a production plan by focusing not on the aggregate level but on every detail associated with each manufactured product. For example, rather than devising a general plan for the aggregate number of TV sets to be sold over the next 12-month period, a detailed production plan focusing on components, subassemblies, and assemblies would be devised for each of the 35 sets in the product line over the same period.

Q13-6. Telephone operating companies face rather difficult production-planning problems partly because they cannot avail themselves of a very useful strategy: inventory.

As an illustration of this consider the manpower-planning problem for inward-outward station movements. Whenever a customer terminates his service or applies for new service, a repairman is sent out to perform the necessary operations. Unfortunately, demand for this service is very cyclical.

Discuss whether or not the smoothing strategies which have been discussed in this chapter could be employed to absorb these fluctuations in demand.

PROBLEMS

✓ **P13-1.** The Mountain Equipment Company manufactures a line of lightweight camping equipment. Demand for their product is quite seasonal, as you can see from next year's forecast.

Month	*Demand*	*Month*	*Demand*
1	10(000) units	7	100(000) units
2	12	8	40
3	15	9	30
4	40	10	20
5	130	11	40
6	200	12	10

(*a*) Given the following production costs and constraints, devise an economical production plan.

Maximum regular-time production/period: 50 units
Maximum overtime production/period: 10 units
Cost of regular-time production: $20/unit
Cost of overtime production: $26/unit
Cost of subcontracting: $29/unit
Inventory carrying cost/period: $2/unit/period
Backordering costs/period: prohibitively high

(*b*) Suppose that an ending inventory of 25 units was required. How would the plan change?

(*c*) What changes would occur if there had been an opening inventory of 25 units and it was required that the ending inventory be 25 units?

✗ (*d*) How would the schedule have changed if regular-time capacity had grown by 1 unit per period?

P13-2. ✓ Given the following demand and production data,

(*a*) Set up problem.

(*b*) Solve, using a computer code.

Period	*Demand*
1	240 units
2	111
3	130
4	60

Initial inventory: 35 units

Regular-time production capacity:
Period 1: 80 units
Period 2: 110
Period 3: 115
Period 4: 120

Overtime production capacity:
Period 1: 15 units
Period 2: 25
Period 3: 30
Period 4: 32

Maximum that can be purchased from subcontractor in each period: 10 units
Closing inventory required: 45 units
Regular-time production costs: $5/unit
Overtime production costs: $7/unit
Inventory carrying costs: $1/unit/period
Backorder costs: $.75/unit/period
Subcontracting costs: $9/unit

P13-3. Using the linear decision rule, determine the work force and production level for the month of January.

Forecast:

Month	*Quantity*	*Month*	*Quantity*
Jan.	400	July	320
Feb.	250	Aug.	350
March	300	Sept.	400
April	350	Oct.	500
May	400	Nov.	600
June	300	Dec.	500

Beginning work force W_{t-1} is 100, and opening inventory is 300 gallons of paint.

P13-4. The administrator for the Smithdale County Hospital is planning his nursing requirements for the next year. The result of a demand forecast is shown below.

Jan.	400(0) man-hours	July	400(0) man-hours
Feb.	420	Aug.	320
March	360	Sept.	300
April	320	Oct.	340
May	530	Nov.	360
June	420	Dec.	410

Presently the maximum available hours per month on regular time is 3,800 man-hours. Overtime can furnish up to 400 additional man-hours per month. Regular-time wages are \$5 per hour while overtime is \$7.50.

In months when the demand for nursing care exceeds the supply, the administrator can utilize a combination of additional volunteers and nurses' aides. When he does this, of course, the quality of service falls. He is, however, bold enough to estimate the cost of this alternative. Including all the imponderables, he feels that the cost per hour of using nonregistered nurses is \$10.

(*a*) What is the schedule which he should follow for the next 12 months?

(*b*) What is the cost of this schedule?

(*c*) Are any of the other smoothing strategies which we have discussed in this chapter relevant to this problem?

(*d*) Suppose that the administrator increases his nursing staff by 160 man-hours per month. That is, he will add one new nurse per month. By December his capacity will therefore be increased by 12 × 160 or 1,920 man-hours. If the additional training costs are \$10,000, is this a better strategy? Assume that new nurses will not work overtime for their first year.

P13-5. Set up the following three-period production-planning problem in linear programming format.

Demand	*Maximum reg.-time production capacity*	*Maximum overtime production capacity*
$D_1 = 40$	$M_1 = 40$	$Y_1 = 5$
$D_2 = 60$	$M_2 = 50$	$Y_2 = 7$
$D_3 = 20$	$M_3 = 35$	$Y_3 = 5$

Overtime cost per unit $v = \$10$
Inventory carrying charge per unit per period $c = \$5$

Cost of hiring in each period:

$h_1 = \$20/\text{unit}$
$h_2 = 21$
$h_3 = 22$

Cost of firing in each period:

$f_1 = \$11/\text{unit}$
$f_2 = 12$
$f_3 = 13$

Regular-time production costs per unit $r = \$9$
Regular-time production output in the previous period $P_0 = 33$

If a standard LP computer code is available, solve the problem.

P13-6. Use the transportation method to set up the following aggregate planning problem:

Regular-time production capacity	10,000 units/period
Overtime production capacity	1,000 units/period
Initial inventory	500 units
Final inventory should be	600 units

Demand:

Period 1	5,000
Period 2	10,000
Period 3	4,000

Regular-time production costs	\$1.30/unit
Overtime production costs	\$2.00/unit
Inventory carrying cost per unit per period	\$1.00

Backorder costs per period are prohibitively high.

If a transportation code is available, solve this problem on a computer.

P13-7. Set up the following three-period aggregate planning problem:

Demand per period $D_i = 50$
Maximum regular-time production:

$M_1 = 50$
$M_2 = 40$
$M_3 = 50$

Maximum overtime production:
$Y_1 = 10$
$Y_2 = 4$
$Y_3 = 5$

Overtime cost per unit $v = \$6$
Inventory carrying cost per unit per period $c = \$2$
Unit cost of increasing production level $h = \$5$
Regular-time production output in the previous period $P_0 = 27$
Unit cost of reducing production level $f = \$3$
Regular-time production cost per period $r = \$4$

If a standard LP code is available, solve the problem.

P13-8. The following is a 12-month demand forecast for the Schmaty Garment Company, a manufacturer of stylish ladies' apparel.

Jan.	150†	May	210	Sept.	200
Feb.	170	June	180	Oct.	220
March	200	July	140	Nov.	210
April	190	Aug.	180	Dec.	200

† All figures are in hundreds of man-hours.

Because of forecast errors which they have grown accustomed to expect, management requires a safety stock level of 10,000 man-hours of merchandise at all times. At present

(December 31) they have 12,000 man-hours' worth of safety stock. In the past, to absorb fluctuations they have shifted their production strategy from a one-shift system to a two-shift system. The capacity on one shift is 18,000 man-hours of output per month, while on a second shift it increases by 5,000 man-hours. Whenever this change is made, a cost of $4,000 is incurred for hiring and training costs. Whenever the shift is eliminated, a cost of $1,000 is incurred. Finally, if inventory is carried, a cost of $0.20 per man-hour per month is incurred.

(*a*) If the only possible strategies for absorbing these fluctuations are inventory and the two production levels, why cannot linear programming techniques be used to solve for the optimal strategy?

(*b*) Employing a strictly intuitive approach, select three alternative strategies and compare their costs. Which one is preferred?

(*c*) What other strategies has the Schmaty Garment Company ignored? Are they likely to be useful in this kind of business?

BIBLIOGRAPHY

Buffa, E. S.: *Production-Inventory Systems: Planning and Control,* Richard D. Irwin, Inc., Homewood, Ill., 1968, chaps. 5, 6, and 7.

Galbraith, Jay R.: "Solving Production Smoothing Problems," *Management Science,* vol. 15, no. 12, August 1969, p. B-665.

Groff, Gene K., and J. F. Muth: *Operations Management: Analysis for Decisions,* Richard D. Irwin, Inc., Homewood, Ill., 1972, chap. 13.

Haehling von Lanzenauer, Christoph: "A Production Scheduling Model by Bivalent Linear Programming," *Management Science,* vol. 17, no. 1, September 1970, p. 105.

Hanssman, F., and S. W. Hess: "A Linear Programming Approach to Production and Employment Scheduling," *Management Technology,* vol. 1, January 1960.

Holt, C. C., F. Modigliani, J. F. Muth, and H. A. Simon: *Planning Production, Inventories, and Work Force,* Prentice-Hall, Inc., Englewood Cliffs, N.J., 1960.

Starr, Martin K.: *Systems Management of Operations,* Prentice-Hall, Inc., Englewood Cliffs, N.J., 1971, chap. 9.

14

The Management and Control of Inventory Systems

INTRODUCTION

The study of inventory problems dates back to 1915, when F. N. Harris developed a very simple but nonetheless useful model of an inventory problem. It was quite natural for the inventory problem to be among the first selected for mathematical analysis, because it was a widespread problem which was seen as being more structured and having a potentially more profitable payoff than the other problems.

Any inventory system requires the collection and processing of large quantities of data. In those pre-computer days the cost of doing this was quite high. Consequently those inventory systems were relatively simple. Starting in the 1950s, however, the computer brought down the cost of handling data. This development sparked renewed interest in inventory theory and marked the beginning of an era in which sophisticated large-scale inventory systems became commonplace in industry.

Traditionally, formal inventory concepts have been applied only to the factory setting. They have been used to control the levels of raw materials, fabricated or finished parts, subassemblies, and finished products. The concept, however, has much broader application. If we define *inventory* as *an idle re-*

source of any kind that has potential economic value, we can transfer the insight which we have gained from our traditional focus. For example, the capacity decision of a hospital can be structured as an inventory problem. A stock of idle beds must be maintained to meet future demand. The consequence of having too many beds is that investment and operating costs are too high but the costs associated with being out-of-stock (no beds available) during periods of peak demand are low since this will seldom happen. On the other hand, with too few beds, investment and operating costs may be low but the out-of-stock costs are high. Presumably there is some number of beds that will minimize the sum of these costs.

As another example we have the stocking of buses in an urban transportation system to meet commuter demand. With too few buses, the investment and operating costs are low but the out-of-stock costs (no seats available) are high. With too many buses, the investment and operating costs are high while out-of-stock costs are low. Again, there is some number of buses which will minimize the sum of these costs.

These problems represent but a few of the ones which will benefit from the insight gained by studying inventory theory. We will start in the next section by developing a framework within which will be placed the subsequent analysis. Then the cost structure of inventory problems is examined, and this is followed by the development of the economic order quantity model. Quantity discounts are then discussed. This is followed by the consideration of stockouts, and then the concept of risk is introduced. Finally, two basic inventory systems are discussed: the fixed order quantity and periodic review systems.

A GENERAL FRAMEWORK FOR INVENTORY MODELS

A source of much confusion to the student who is confronted with inventory theory for the first time is the plethora of inventory models. In an attempt to minimize this confusion, a general framework will be developed within which these models will eventually be placed.

DEMAND

First we will consider demand. According to our definition, inventory decisions are made with reference to a future demand. Three states of knowledge can exist with regard to this level. It may be possible that we know with *certainty* what the future level of demand will be. For example, in some construction projects, the cement, steel, glass, and lumber requirements are known with certainty well in advance. The usual case, however, is that we are able only to estimate the probability distribution of the future level of demand. By utilizing sales records, for example, we could determine the likelihood that several different levels might exist. One such distribution is shown in Fig. 14-1. It can be seen that there is a 45 percent chance that demand will be 11,000 during any week, a 23 percent chance that it will be 10,000, and so on. In this case we have an inventory problem under *risk.*

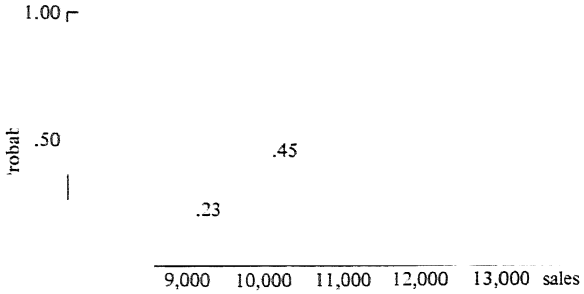


Fig. 14-1 Probability distribution of demand.

In the last case we know nothing about the likelihood of future levels of demand. This is called an inventory problem under *uncertainty,* and it is usually as uncommon as the case of certainty. It is highly unlikely that we know absolutely nothing about demand.

LEAD TIME

The second aspect of this framework is the knowledge we have of the lead time. The *lead time* is the period between the time when an order is placed and the time when it is received. On occasion we know this lead time with certainty. For example, there might be a particular item in inventory which has always taken exactly 7 working days from initiation of order to delivery in stock. Again, the more likely case is that lead time can best be described by a probability distribution. This is the case of lead time under risk. Finally we have the situation of lead time under uncertainty—the case where we are totally ignorant of the lead time. This state of knowledge is again quite unlikely.

INSIDE OR OUTSIDE PROCUREMENT

The third aspect has to do with whether the item under consideration will be purchased from outside the firm or will be produced within it. When the item is produced within the firm, it affects the production process. In the chapter on aggregate production planning we saw, quite explicitly, the interaction between inventory and production. When we established production plans, inventory decisions were but one of the many factors that were considered. Imagine the consequences if an unusually large set of orders were released to the shop from the inventory control department. If little interaction had taken place between the shop and inventory control, the shop might respond by hiring, overtime, and, possibly, multiple shifts. Next month, however, the number of orders released might be below average. The shop's response might include firing and undertime. Certainly a production-smoothing approach would be much more effective. Perhaps order size could have been reduced in the first month and the balance run in the second month. Several solutions exist; however, the point is that inventory decisions for parts produced within the firm are more complex than similar decisions for those parts when procured from outside suppliers.

STATIC AND DYNAMIC PROBLEMS

The fourth aspect has to do with whether the inventory decision is made only once or whether it is continually made. For example, the inventory decision for Christmas trees is made only once. An order is placed in the fall and there is usually no opportunity for replenishment orders should demand exceed the supply. If, on the other hand, supply exceeds demand, the leftovers must be scrapped. This *static* case is quite unlike the more common case where an order is but one in a sequence of orders through time. For example, a drug in a hospital pharmacy is continually reordered. If an inadequate supply was ordered, the next order would be placed sooner than usual; or if demand for the drug had fallen off, the next order would be later than usual. In this *dynamic* situation, one decision directly affects the next, and it is necessary to reflect this situation in the analysis.

BEHAVIOR OF DEMAND THROUGH TIME

Another aspect which must be included in our framework is the behavior of demand through time. The first case is when demand, or the probability distribution of demand, does not change over time. For example, demand would be 10 today, 10 the next day, 10 the day after, and so on. The other case is when demand or the probability distribution of demand varies over time. An example of this would be a demand level of 10 today, 15 tomorrow, 13 the next day, and so on. If the probability distribution varied, one would have one relevant probability distribution of demand for today, another one for tomorrow, and so on.

COSTS

A major aspect of this framework is the cost structure of the problem. In some cases price discounts are relevant, while in others they are not. Stockouts may be acceptable in one situation and quite unacceptable in another. The next section takes a close look at these and other relevant costs.

INVENTORY COSTS

The solution of any inventory problem requires the answers to two questions: First, *when* or how often should an order be placed? Second, *how much* should be ordered? If we set as our objective the minimization of costs, our first step should be to identify the relevant costs. From this cost model we will be able to determine the answers to these two questions.

PER-PIECE PROCUREMENT COSTS

The first major cost is the per-piece procurement cost. In the simple case the per-piece price remains the same regardless of quantity purchased, but in the more complex case a series of price breaks is offered for purchases of increasing quantity.

ORDERING COSTS

A second kind of procurement cost is the *fixed cost* associated with placing an order. This includes the cost of typing the purchase order, expediting the order, placing the item in stock, and so on. All costs that do not vary with the size of the order are included in this category. When the item is produced within the firm rather than ordered from an outside source, setup costs are also included.

INVENTORY CARRYING COSTS

To carry inventory costs money. This cost can be broken down into several components. The first is the opportunity cost of money tied up in inventory.

The funds that are allocated to inventory could be used for other productive purposes. Since the capital-budgeting process ensures that the expected earnings of other projects will be at least as great as the firm's cost of capital, it seems reasonable to consider that figure as the opportunity cost of money tied up in inventory.

To this opportunity cost we must also add storage, deterioration, and insurance costs. If additional storage space must be built or rented, it becomes a legitimate marginal expense. If, on the other hand, there is plenty of space in the warehouse and the idle space cannot be put to any productive use, there are no marginal storage costs.

Many kinds of goods suffer from some kind of deterioration or obsolescence during storage. For example, dry cell batteries and vegetables deteriorate, while electronic equipment and fashion wear slowly become obsolete. When building a cost model of an inventory system, this is a very real cost which must be included. It will serve to discourage large-quantity, infrequent orders for goods which suffer from this affliction.

The last two costs associated with carrying an inventory are insurance costs and inventory taxes. Now we can summarize the cost of carrying inventory in the following way:

$$I = C + G + K + T$$

where I = cost of carrying inventory in percent per period
C = opportunity cost of carrying inventory in percent per period
G = storage costs in percent per period
K = deterioration and obsolescence costs in percent per period
T = insurance and taxes in percent per period

STOCKOUT COSTS

Running out of stock also costs money. This cost is called a *stockout* cost and can lead to one of two possible outcomes. The first is that this out-of-stock position will lead to a backorder. This situation is frequently found in mail-order houses where an item is temporarily out-of-stock and the customer is willing to wait until it is replenished. The firm does suffer some cost in this situation, since at the very least some expediting and extra communication, perhaps in the form of telephone calls or postcards, with the customer must take place. In addition to the explicit costs, repeated backorders will certainly lead to an erosion of goodwill. This effect, however, is very difficult to measure.

The second possible outcome of an out-of-stock position is the lost sale. Here the cost is much more severe, but the exact level of the cost is, once again, very difficult to measure. In this case, the customer places an order, receives an out-of-stock response, and takes his business elsewhere. The cost of this occurrence to the firm is more than the profit lost on the potential sale. The company must take into account the likelihood that the customer will not return and that therefore the profit on future sales might also be lost. In addition, the loss in goodwill that this will precipitate may also persuade others to purchase elsewhere.

Since backorder and lost sales costs are hard to estimate, service levels are frequently specified. For example, management may feel that stockouts should not occur more than 2 percent of the time. Unconsciously, however, management is imputing a cost to stockouts. If they had said that there should never be a stockout, they would be implying that there is an infinite cost associated with one. A policy of "no stockouts" is therefore very suspicious.

This completes the development of our structure; as we progress from simple inventory models to more complex ones, their place in this theoretical structure will be carefully pointed out. The framework, however, is not exhaustive, yet it should adequately serve to unify our introduction to inventory theory.

EOQ MODEL

The first inventory problem which we will model is a very simple one. First we assume that demand and lead time are known with certainty. In addition we consider that the item will be purchased from outside the firm and that demand will continue well into the future. We also assume not only that demand is known with certainty, but that it is the same from day to day and that stockouts are not allowed. Under these assumptions, Fig. 14-2 depicts the inventory position through time.

If we start to observe the inventory position immediately after receipt of an order, the quantity in stock Q decreases steadily until a lead time's supply is reached. If lead time is 4 days and demand is 3 per day, a lead time's supply is 12 units. Therefore, when 12 units remain, an order is placed. This

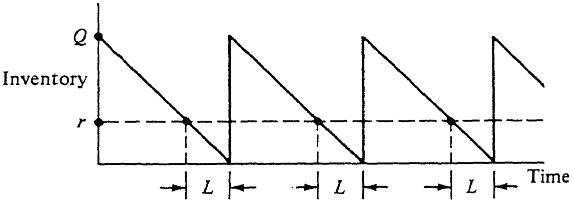


Fig. 14-2 Time pattern of inventory level for simple EOQ model.

is called the *reorder point r*. Exactly 4 days after the order is placed, the stock is replenished and the cycle repeats itself.

OPPOSING COSTS

If the quantity ordered each time is large, the orders will be placed infrequently. As a result the average inventory will be high, and this will lead to high inventory carrying costs. With these infrequent large orders, however, the fixed cost of ordering incurred each year will be small.

Let us look at the other extreme, where small orders are placed very frequently. The average inventory in this case will be small and the inventory carrying costs will therefore be low. Since orders will be placed frequently, the fixed ordering costs incurred each year will be high. What becomes apparent, then, is that there are costs which move in opposite directions. This is illustrated in Fig. 14-3. We see that when order quantities increase in size, carrying costs increase because average inventory size becomes larger. Fixed order costs, on the other hand, decrease as order quantities increase because fewer orders are placed per year.

Given that the minimization of total costs is the objective, the carrying costs and fixed order costs are added together. The result is the total-cost curve shown in Fig. 14-3. The best strategy is to order Q^* each time an order is placed. If any other quantity is ordered, a higher total cost will be incurred. Next we turn to formulating this problem as a mathematical model.

TOTAL COSTS

The total cost of any order quantity Q can be written in the following way:

$$\begin{pmatrix}\text{Total cost}\\ \text{per period}\end{pmatrix} = \begin{pmatrix}\text{fixed order costs}\\ \text{per period}\end{pmatrix} + \begin{pmatrix}\text{inventory carrying}\\ \text{costs per period}\end{pmatrix}$$

If we let A = fixed order costs per order
D = demand per period (usually one year)
C = per-piece procurement cost
I = cost of carrying inventory in percent per period

the fixed order cost per period can be written as $(D/Q)A$, where D/Q is the number of orders placed per period.

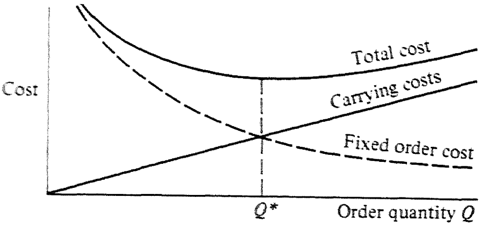


Fig. 14-3 Inventory carrying costs.

The inventory carrying costs can be expressed in the following way:

$$\begin{pmatrix}\text{Inventory}\\ \text{carrying}\\ \text{cost per}\\ \text{period}\end{pmatrix} = \begin{pmatrix}\text{average}\\ \text{inventory}\end{pmatrix} \times \begin{pmatrix}\text{per-piece}\\ \text{procurement}\\ \text{cost}\end{pmatrix} \times \begin{matrix}\text{inventory}\\ \text{carrying cost}\\ \text{in percent per}\\ \text{period}\end{matrix}$$

The average inventory during any cycle is

$$\text{Maximum inventory} + \text{minimum inventory}$$

and since each cycle is just like any other, this also represents the average inventory per period. In our example the minimum inventory is zero and the maximum inventory is equal to the order quantity Q. We can, therefore, write the inventory carrying cost per period as $(Q/2)\ (C)\ (I)$.

Combining the fixed order and inventory carrying costs, we have

$$\text{TC} = \frac{D}{Q} A + \frac{Q}{2} CI$$

This is the total-cost expression for any quantity Q. Since the decision maker has control over the quantity ordered, the relevant question is, What order quantity Q will minimize total costs TC?

MINIMUM COST

By resorting to differential calculus, we can determine the value of Q at which this total-cost function will be minimized. This procedure is illustrated in the next optional section, and from it we learn that if the following quantity is ordered, our objective will be achieved:

$$Q^* = \sqrt{\frac{2DA}{CI}}$$

This quantity is frequently called the *economic order quantity* formula or EOQ. The total cost of this strategy is

$$\text{TC} = \frac{D}{Q^*} A + \frac{Q^*}{2} CI$$

$$\text{TC} = \frac{D}{\sqrt{2DA/IC}} A + \sqrt{2DA/IC}$$

which can be simplified to

$$\text{TC} = \sqrt{2DAIC}$$

Example Consider the following example. The yearly demand for a particular item has been forecast as 10,000 units. The per-piece procurement cost is \$1, the fixed order costs are \$10 per order, and the inventory carrying costs determined to be 20 percent per unit per year. The EOQ can be computed as

$$D = 10{,}000 \text{ units}$$
$$A = \$10 \text{ per order}$$
$$C = \$1 \text{ per unit}$$
$$I = 20\%$$
$$Q^* = \sqrt{2DA}$$
$$Q^* = \sqrt{\frac{2(10{,}000)(10)}{.20(1)}} = 1{,}000 \text{ units}$$

Therefore 1,000 units will be ordered at a time, and the total cost of this strategy is

$$\text{TC} = \sqrt{2DAIC} + \sqrt{2(10{,}000)(10)(.20)(1)} = \$200$$

//// DERIVATION OF EOQ

Since the total-cost expression is a function in one variable Q, the minimum can be determined by differentiating the function, setting it equal to zero, and solving for Q. This is accomplished in the following way:

$$\text{TC} = \frac{D}{Q}A + \frac{Q}{2}CI$$
$$\frac{d(TC)}{dQ} = -\frac{DA}{Q^2} + \frac{CI}{2}$$
$$-\frac{DA}{Q^2} + \frac{CI}{2} = 0$$
$$Q^* = \sqrt{\frac{2DA}{IC}}$$

////

SENSITIVITY ANALYSIS

Suppose in the example given earlier that the actual cost of processing an order was \$20 and that we had erroneously measured it as \$10. Therefore the true EOQ should have been

$$Q^* = \sqrt{\frac{2(10{,}000)(20)}{.20(1)}} = 1{,}414$$

and the total cost of the strategy that should have been adopted is

$$\text{TC} = \frac{10{,}000}{1{,}414}20 + \frac{1414}{2}(1)(.20) = \$282.80$$

However, the total cost of the strategy that *was* adopted actually is

$$\text{TC} = \frac{10{,}000}{1{,}000}(20) + \frac{1{,}000}{2}(1)(.20) = \$300$$

Therefore with a 100 percent error in the fixed order cost, the total costs were higher by

$$\frac{300 - 282.80}{282.80} = 0.064 \qquad \text{or} \qquad 6.4\%$$

We can conclude that the analysis is not very sensitive to the estimation of the relevant costs.

//// A MORE FORMAL APPROACH TO SENSITIVITY ANALYSIS

Suppose that the actual carrying costs and fixed order costs are I and A. If, however, they are measured as k_1I and k_2A, the computed EOQ would be

.20 .5

$$Q^* = \sqrt{\frac{2Dk_2A}{k_1IC}}$$

The actual cost incurred with this erroneous EOQ would be

$$\text{TC} = \frac{D}{Q^*}A + \frac{Q^*}{2}IC$$

$$\text{TC} = \frac{D}{\sqrt{(2Dk_2A/k_1IC)}}A + \frac{\sqrt{(2Dk_2A/k_1IC)}}{2}IC$$

This can be simplified to:

$$\text{TC} = \left(\sqrt{\frac{k_1}{k_2}} + \sqrt{\frac{k_2}{k_1}}\right)\sqrt{DAIC}$$

Had the true inventory and fixed order costs been used, the total cost would have been

$$\text{TC} = \sqrt{2DAIC}$$

The percentage error can be determined in the following way:

$$\frac{\text{Actual TC} - \text{true optimal TC}}{\text{True optimal TC}}$$

$$\frac{(\sqrt{k_1/k_2} + \sqrt{k_2/k_1})\sqrt{DAIC/2} - \sqrt{2DAIC}}{\sqrt{2DAIC}}$$

This can be simplified to

$$\frac{1}{2}\left(\sqrt{\frac{k_1}{k_2}} + \sqrt{\frac{k_2}{k_1}}\right) - 1$$

Consider a case where inventory carrying costs have been overestimated 100 percent. We have $k_1 = 2$. In addition fixed order costs have been underestimated by 20 percent, or $k_2 = .8$. Using these erroneous values to compute EOQ, we will be off from the optimal by

$$\frac{1}{2}\left(\sqrt{\frac{2}{.8}} + \sqrt{\frac{.8}{2}}\right) - 1 = 0.11 \qquad \text{or} \qquad 11\%$$

Thus we see that the error in total cost depends upon the error in each of these parameters and, furthermore, that large errors have only a modest influence on total costs. That is, total costs are generally insensitive to cost estimates. ////

QUANTITY DISCOUNTS

The EOQ model assumes that one price prevails for all quantities purchased. Frequently, however, price breaks are offered for larger orders. Consider the following price schedule:

If $Q < q_0$, the price is C_1

If $Q \geq q_0$, the price is C_2

where Q = quantity ordered

q_0 = price break quantity

This being the case, the total procurement cost is no longer constant and must be included in the total-cost expression. We now have

$$\text{TC} = CD + \frac{D}{Q}A + \frac{Q}{2}IC$$

The decision process in the presence of this price break is quite complex, and it is helpful to develop the analysis as a sequence of decisions.

The first step is to determine the economic order quantity Q^* based on the higher price C_1. This is recorded as step 1 in Fig. 14-4. If Q^* is greater than the price-break point q_0, the optimal strategy is to order the economic order quantity based on the lower price C_2.

$$Q^{**} = \sqrt{\frac{2DA}{IC_2}}$$

The reason behind this strategy can be seen from Fig. 14-5(a). Whenever Q^* is greater than q_0, the total-cost curve associated with the lower price lies below the total-cost curve for the higher price in the region beyond q_0. Therefore, in the interest of cost minimization, Q^{**} is ordered.

If Q^* is not greater than the price break q_0, the economic order quantity at the lower price must be computed and compared with q_0. If Q^{**} is greater than q_0, the cost functions must look like Fig. 14-5(b). If you recall, we have

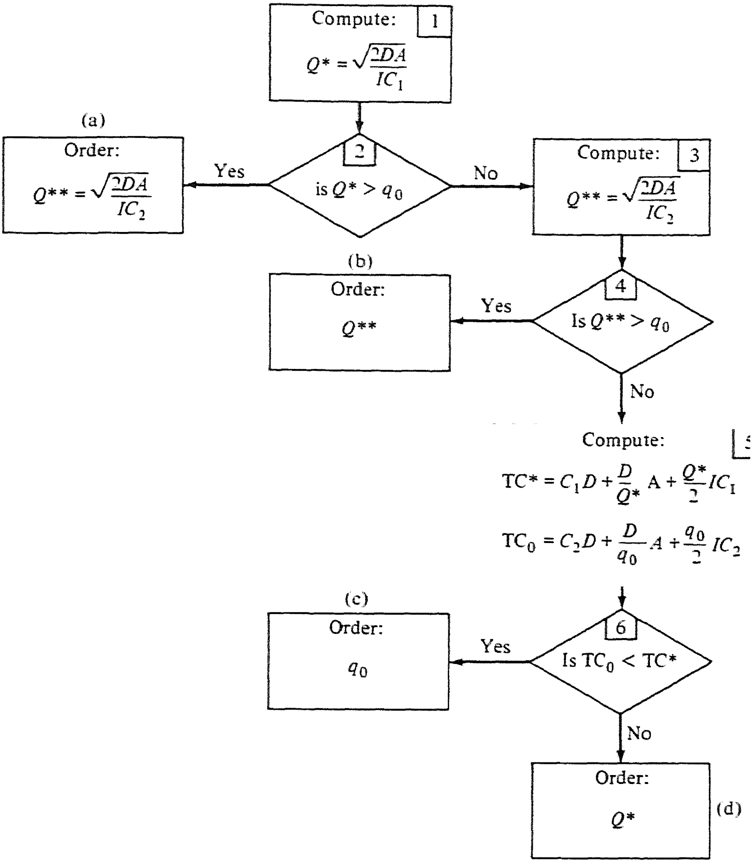


Fig. 14-4 Quantity discount decision model.

already determined that Q^* lies below q_0, and now we have established that Q^{**} lies above q_0. Given this outcome, the optimal strategy is to order Q^{**}.

If Q^{**} is not greater than q_0, the situation depicted in either Fig. 14-5(*c*) or Fig. 14-5(*d*) prevails. Both Q^* and Q^{**} must be less than q_0, and the only reasonable comparison is between the total cost associated with Q^* and the total cost incurred if exactly q_0 were ordered. Said another way, the decision is either to order the economic quantity based on the lower price or to order just enough to get the price break. Therefore the total costs for each alternative must be computed. First, the total cost incurred at Q^* is computed:

$$\text{TC}^* = C_1D + \frac{D}{Q^*}A + \frac{Q^*}{2}IC_1$$

Then the total cost if just the price-break quantity were purchased is computed.

$$\text{TC}_0 = C_2D + \frac{D}{q_0}A + \frac{q_0}{2}IC_2$$

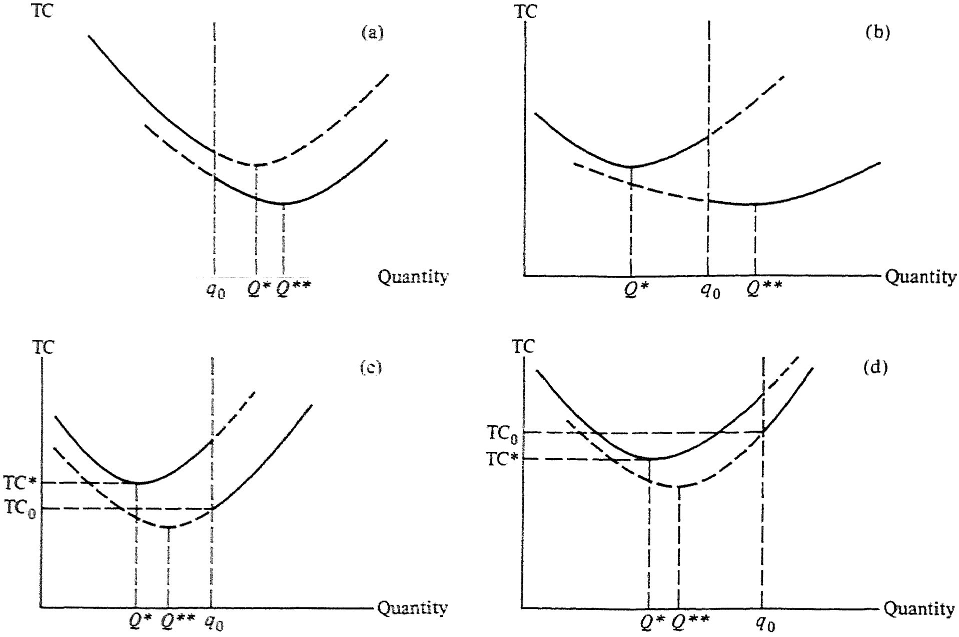


Fig. 14-5 Cost functions for various discount situations.

If TC* is less than TC_0, Q^* is ordered and the behavior of the cost functions is as shown for those of Fig. 14-5(*d*). If TC_0 is less than TC*, the relevant cost functions are shown in Fig. 14-5(*c*), and q_0 should be ordered.

Example Consider the following:

D = 10,000 units If $Q < q_0$, $C_1 = \$1.00$
A = \$10 per order If $Q \geq q_0$, $C_2 = \$.80$
I = 20%
q_0 = 1,700 units

Step 1: $$Q^* = \sqrt{\frac{2(10{,}000)(10)}{.20(1)}} = 1{,}000$$

Step 2: $$Q^* \not> 1{,}700$$

Step 3: $$Q^{**} = \sqrt{\frac{2(10{,}000)(10)}{.20(.80)}} = 1{,}580$$

Step 4: $$Q^{**} \not> q_0$$

Step 5: $$TC^* = 1(10{,}000) + \frac{10{,}000}{1{,}000}10 + \frac{1{,}000}{2}(.20)(1)$$

$$TC^* = 10{,}000 + 100 + 100$$

$$TC^* = 10{,}200$$

$$TC_0 = .80(10{,}000) + \frac{10{,}000}{1{,}700}\,10 + \frac{1{,}700}{2}(.20)(.80)$$

$$TC_0 = 8{,}000 + 59 + 136$$

$$TC_0 = 8{,}195$$

Step 6: $TC_0 < TC^*$

Therefore,

Order $q_0 = 1{,}700$

EOQ MODEL WITH STOCKOUTS ALLOWED

Now we will return to our simple EOQ model, but this time we will relax the assumption that no stockouts are allowed. Again the first step is to specify the total-cost expression.

TOTAL COSTS

Writing the total-cost expression for those costs which are affected by the order quantity Q, we have

$$TC = \begin{pmatrix}\text{fixed order}\\ \text{costs per}\\ \text{period}\end{pmatrix} + \begin{pmatrix}\text{inventory}\\ \text{carrying}\\ \text{costs per period}\end{pmatrix} + \begin{pmatrix}\text{backorder}\\ \text{costs per}\\ \text{period}\end{pmatrix}$$

The fixed order costs per period are computed just as before:

$$\begin{pmatrix}\text{Fixed order}\\ \text{costs per}\\ \text{period}\end{pmatrix} = \frac{D}{Q}A$$

The other costs, however, require that we look at the time pattern of inventory levels. This is shown in Fig. 14-6.

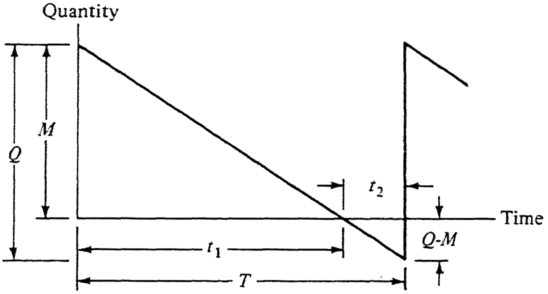


Fig. 14-6 Time pattern of inventory levels when stockouts are incurred.

Turning to the inventory carrying cost per period, we have the following:

$$\begin{pmatrix}\text{Inventory}\\ \text{carrying cost}\\ \text{per period}\end{pmatrix} = \begin{pmatrix}\text{average}\\ \text{inventory}\end{pmatrix} \times \begin{pmatrix}\text{per-piece}\\ \text{procurement}\\ \text{cost}\end{pmatrix} \times \begin{pmatrix}\text{inventory}\\ \text{carrying cost}\\ \text{in \% per period}\end{pmatrix}$$

$$\begin{pmatrix}\text{Inventory}\\ \text{carrying cost}\\ \text{per period}\end{pmatrix} = \begin{pmatrix}\text{average}\\ \text{inventory}\end{pmatrix} (C)(I)$$

Average inventory, however, is quite different from what it was in the simple case. The maximum inventory is M and the minimum is zero. The average inventory of $(M + 0)/2 = M/2$, however, is carried for only t_1/T of the cycle. The average inventory for the remainder of the cycle is zero. We can therefore write the average inventory as $(M/2)\ (t_1/T)$. Therefore,

$$\begin{pmatrix}\text{Inventory carrying}\\ \text{cost per period}\end{pmatrix} = \frac{M}{2}\left(\frac{t_1}{T}\right)(C)(I)$$

Next we turn to the backorder costs per period. These costs can be determined in the following way.

$$\begin{pmatrix}\text{Backorder}\\ \text{costs per}\\ \text{period}\end{pmatrix} = \begin{pmatrix}\text{average number}\\ \text{of backorders}\\ \text{per period}\end{pmatrix} \times \begin{pmatrix}\text{backorder}\\ \text{cost}\\ \text{per period}\end{pmatrix}$$

The average number of backorders *while* backorders are taking place is equal to the maximum number of backorders divided by 2, or $(Q - M)/2$. Since backorders occur only during (t_2/T) of the cycle, the average number of backorders per cycle is $[(Q - M)/2]\ (t_2/\mathrm{T})$. We therefore have

$$\begin{pmatrix}\text{Backorder}\\ \text{costs per}\\ \text{period}\end{pmatrix} = \frac{Q - M}{2}\,\frac{t_2}{T}\,B$$

where B = cost incurred for every backorder

Returning to the expression for total cost, we now have

$$\mathrm{TC} = \frac{D}{Q}A + \frac{M}{2}\,\frac{t_1}{T}\,CI + \frac{Q - M}{2}\,\frac{t_2}{T}\,B$$

We can, however, remove t_1, t_2, and T from this expression in the following way:

$$t_1 = \frac{M}{D}$$

$$T = \frac{Q}{D}$$

$$t_2 = T - t_1 = \frac{Q}{D} - \frac{M}{D} = \frac{Q - M}{D}$$

Therefore,

$$\frac{t_1}{T} = \frac{M}{Q}$$

$$\frac{t_2}{T} = \frac{Q - M}{Q}$$

Returning to our total-cost expression, we now have

$$\mathrm{TC} = \frac{D}{Q}A + \frac{M}{2}\left(\frac{M}{Q}\right)CI + \frac{Q - M}{2}\left(\frac{Q - M}{Q}\right)B$$

$$\mathrm{TC} = \frac{D}{Q}A + \frac{M^2}{2Q}CI + \frac{(Q - M)^2}{2Q}B$$

The next step is to find that order strategy Q which minimizes these costs.

MINIMUM COST

Since we now have a function in two variables Q and M, it becomes a little more difficult to solve for the lowest total cost TC*. In the next *optional* section the following solution is found.

$$Q^* = \sqrt{\frac{2DA}{IC}}\sqrt{\frac{CI + B}{B}}$$

and $$M = \frac{QB}{CI + B}$$

Notice that if backorder costs B are very large, the value of the second radical approaches 1 and economic order quantity derived for the simple case holds.

Example Consider the following:

D = 10,000 units

A = \$10 per order

I = 20%

C = \$1 per unit

B = \$5 per each backorder incurred

$$Q^* = \sqrt{\frac{2(10{,}000)(10)}{.20(1)}}\sqrt{\frac{1(.20) + 5}{5}}$$

$$Q^* = 1{,}000(1.02)$$

$$Q^* = 1{,}020$$

$$M = \frac{1{,}020(5)}{1(.20) + 5} = 980 \text{ units}$$

Therefore the maximum amount backordered is

$$Q - M = 1{,}020 - 980 = 40 \text{ units}$$

//// **DERIVATION OF ECONOMIC ORDER QUANTITY WITH BACKORDERS ALLOWED**

The process by which the optimal value of Q is found involves taking the partial derivative of the total-cost expression first with respect to the variable Q and then with respect to the variable M. Then the partial derivatives are set equal to zero and solved for Q.

The total-cost expression is repeated for convenience:

$$\mathrm{TC} = \frac{D}{Q}A + \frac{M^2}{2Q}CI + \frac{(Q - M)^2}{2Q}B$$

Its partial derivative with respect to M is

$$\frac{\partial(\mathrm{TC})}{\partial M} = \frac{M}{Q}CI - \frac{(Q - M)}{Q}B$$

Then we set it equal to zero,

$$\frac{M}{Q}CI - \frac{(Q - M)}{Q}B = 0$$

and solve for M:

$$M = \frac{QB}{CI + B}$$

Next we take the partial derivative of the total-cost expression with respect to Q.

$$\frac{\partial(\mathrm{TC})}{\partial Q} = -\frac{DA}{Q^2} - \frac{M^2}{2Q^2}CI + \frac{B}{2} - \frac{M^2B}{2Q^2}$$

Setting it equal to zero, we have the following:

$$-\frac{DA}{Q^2} - \frac{M^2}{2Q^2}CI + \frac{B}{2} - \frac{M^2B}{2Q^2} = 0$$

Finally, by substituting M in this equation and simplifying, we get the expression for Q^*:

$$Q^* = \sqrt{\frac{2DA}{IC}}\sqrt{\frac{CI + B}{B}}$$

////

INVENTORY MODELS UNDER RISK

In the simple EOQ model it was assumed that demand and lead time were both known with certainty. Under this condition, whenever inventory levels reached a lead time's worth of demand, an order for Q^* was placed. Then as the stock was finally depleted, a replenishment order would arrive. We did find, however, that if backorders had a finite cost, a reorder level below the average lead time demand would be appropriate. This strategy would result in some desired number of backorders. We can, therefore, conclude that the purpose of reorder levels is not to prevent stockouts, but to keep them within desired limits.

In the case of fluctuating demand, it may be quite logical to set reorder levels above the average lead time demand. For if we reorder when there is only enough in stock to meet the average demand during lead time, an out-of-stock position will occur whenever demand exceeds this average value. In fact, the average will be exceeded half of the time. To protect from undesirably large stockout situations, safety stocks are maintained. They will provide the cushion needed whenever demand exceeds this average.

Safety stocks can be defined as the difference between the reorder level r and the average lead time demand $\bar{u}$.

$$S = r - \bar{u}$$

Therefore as the reorder point is raised, the safety stock increases and the likelihood of a stockout during any cycle decreases.

The inventory pattern through time for a particular item which exhibits fluctuating demand is shown in Fig. 14-7. An order is placed whenever stock levels fall to r. On the average, stock will be depleted to S when the replenishment order arrives. During the first cycle we see that demand was about average; however, in the second cycle, the demand was much greater than average. Protecting the system from a stockout during the lead time was a safety stock maintained at S. If more protection was desired, the level of r could have been increased. With a higher r, the safety stock would have been larger and the chances of a stockout occurring in any one cycle would have been lower.

SETTING THE REORDER POINT WHEN THE SERVICE LEVEL IS GIVEN

Consider the distribution of demand during lead time shown in Table 14-1. The data from which this table was compiled were collected over several periods. In 2 percent of these periods demand was 44; in 3 percent of them, demand was 45; and so on. The last column in Table 14-1 illustrates the probability that

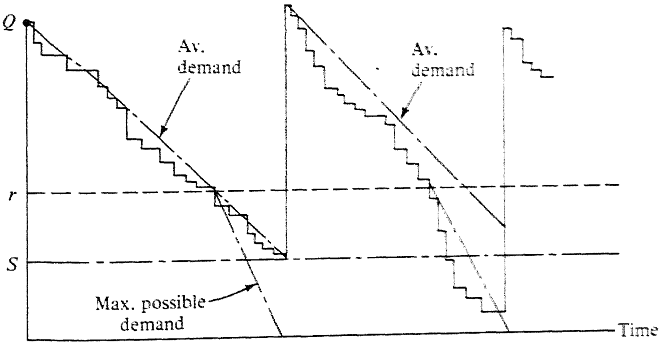


Fig. 14-7 Inventory levels under conditions of fluctuating demand.

Table 14-1 Lead time demand

Lead time demand u	$p(u)$	*Probability that demand will exceed u*	u	$p(u)$	*Probability that demand will exceed u*
			50	.10	.54
			51	.11	.43
43	.00	1.00	52	.10	.33
44	.02	.98	53	.09	.24
45	.03	.95	54	.08	.16
46	.04	.91	55	.07	.09
47	.08	.83	56	.05	.04
48	.09	.74	57	.03	.01
49	.10	.64	58	.01	.00

demand will exceed any of these levels, u. For example, the probability that demand will exceed 56 units is .04. The last column will be very useful for setting safety stocks when a service criterion is given.

In many cases, it might be quite difficult to determine the actual cost of a stockout. An indirect way to assess this is to specify a service level. For example, the manager of the product line whose lead time was recorded in Table 14-1 might feel that the likelihood of a stockout should not exceed 4 percent during any one cycle. Therefore, the reorder point should be set at $r = 56$. We would, therefore, expect an out-of-stock situation to occur in only one out of 25 ordering periods. This implies that the cost of a stockout for this item is quite high.

Suppose that the lead time distribution recorded in Table 14-1 could be approximated by a normal distribution as shown in Fig. 14-8. The average level of demand is shown as u and the standard deviation is given as σ. If the reorder level is set at r, the safety stock is S and the probability that demand will exceed the reorder level r is shown as the shaded area.

Consider an example where $u = 400$ and $\sigma = 15$. Can you find the appropriate safety stock level such that the likelihood of a stockout during any one cycle is less than 1 percent? From the normal table we can see that r should be set 2.33 standard deviations from the mean $\bar{u}$.

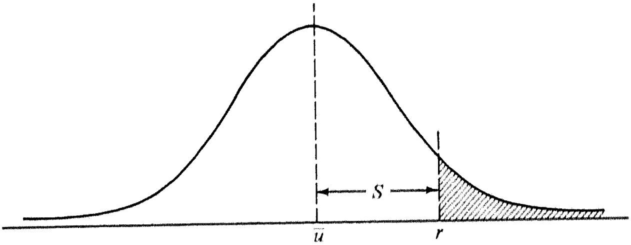


Fig. 14-8 Distribution of lead time demand.

$$r = \bar{u} + 2.33(\sigma)$$
$$r = 400 + 2.33(15)$$
$$r = 435$$

Therefore, the reorder point is set at 435 and the safety stock is $435 - 400 = 35$ units.

Now that we have examined a way in which to set safety stock levels when the service criterion is specified, we turn to the case where we must set safety stock levels when backorder costs are known.

SETTING THE REORDER POINT WHEN BACKORDER COSTS ARE GIVEN

The total cost when backorders are allowed can be written as

$$\text{TC} = \frac{D}{Q}A + \frac{Q}{2}CI + (r - \bar{u})CI + \frac{D}{Q}[E(N)]B \tag{14-1}$$

where the first term represents the fixed order costs and the second represents the carrying costs for all but safety stocks. The third term represents the carrying costs associated with safety stock. In the last term,

$\frac{D}{Q}$ = number of orders placed per period
$E(N)$ = expected number of backorders per cycle
B = cost incurred per backorder

Their product represents the expected cost of backorders per period.

Solution of this cost function for the *optimal* values of Q and r (Q^* and r^*) is quite complex and beyond the scope of this chapter. There is, however, an approach which does not provide the theoretically optimal solution but does furnish a surprisingly good one. Its benefit is that it makes the process of finding a solution for Q^* and r^* relatively simple. It is to this method that we now turn.

The first step is to isolate the first two terms of the total-cost expression and solve for Q^*. Since these two terms are precisely those found in our simplest inventory model, the EOQ formula can be used to determine Q^*.

The second step is to use this value of Q^* in optimizing the last two terms of the total-cost equation with respect to r. To illustrate these steps, we now turn to an example.

The data for a particular unit kept in stock are

D = 1,000 units per year
C = \$10 per unit
A = \$40 per order
I = 20% per year
B = \$5 per unit backlogged

and Table 14-1 represents the probability distribution of demand during lead time.

First, Q^* is computed:

$$Q^* = \sqrt{\frac{2(1{,}000)(40)}{.20(10)}} = 200$$

Next the last two terms in the total-cost expression must be solved for r^*. The first of these terms, $(r - \bar{u})CI$, is the cost of carrying the safety stock, and you would expect this to *increase* as safety stocks are increased. The last term represents the expected cost of backorders per period; it *decreases* as safety stocks are increased. These last two terms, then, move in opposite directions, and the problem, therefore, is to find that level of safety stock which will minimize the sum of these costs.

From Table 14-1 the average lead time demand can be computed in the following way:

$$\bar{u} = \sum_{i=1}^{n} u_i[p(u_i)]$$

$$\bar{u} = 50$$

In Table 14-2 we will determine the consequence of reordering at this average lead time demand level and above.

Consider a reorder level set at $r = 52$. The probability of demand during lead time equaling this is $p(u) = .10$. This is entered in row 1 of Table 14-2. Since the average demand was computed to be $\bar{u} = 50$, the safety stock for a reorder level of 52 would be 2. This is entered in row 2. The annual carrying cost associated with this size safety stock can then be computed.

$$(r - \bar{u})CI$$

$$2(10)(.20) = \$4$$

This is entered in row 3. Next we turn to the computation of backorder costs.

If the reorder level is set at $r = 52$, a backorder will occur whenever lead time demand exceeds this level. The expected number of backorders per cycle for each reorder level r can be computed in the following way.

$$E(N) = \sum_{u>r} (u - r)p(u)$$

In Table 14-3, this procedure is illustrated for a reorder level of $r = 52$. In this case, the expected number of backorders per cycle is .87. This is entered in row 4 of Table 14-2. Now we can compute the backorder cost for this strategy and enter it in row 5:

$$\frac{D}{Q^*}[E(N)]B$$

$$\frac{1{,}000}{200}(.87)(5) = 21.75$$

Table 14-2 Sum of safety stock carrying costs and backorder costs

		Reorder level r								
		50	51	52	53	54	55	56	57	58
(1)	Probability of demand $p(u)$	.10	.11	.10	.09	.08	.07	.05	.03	.01
(2)	Safety stock $(r - \bar{u})$	0	1	2	3	4	5	6	7	8
(3)	Annual carrying cost $(r - \bar{u})CI$	0	\$ 2	\$ 4	\$ 6	\$ 8	\$10	\$12	\$14	\$16
(4)	Expected number of backorders $E(N) = \sum_{u>r} (u - r)p(u)$	1.84	1.30	.87	.54	.30	.14	.05	.01	.00
(5)	Backorder cost $\frac{D}{Q^*}[E(N)]B$	\$46.00	\$32.50	\$21.75	\$13.50	\$ 7.50	\$ 3.50	\$ 1.25	\$ .25	\$ .00
(6)	Total cost $(r - \bar{u})CI + \frac{D}{Q^*}[E(N)]B$	\$46.00	\$34.50	\$25.75	\$19.50	\$15.50	\$13.50	\$13.25	\$14.25	\$16.00

Table 14-3 Computation of expected number of backorders

If $r = 52$ and actual demand is	*The backlog would be $(u - r)$*	*The probability of this is $p(u)$*	*Computation of the expected number of backorders per cycle $(u - r)p(u)$*
$\underline{\angle}$ 52	0	.67	$0 \times .67 = .00$
53	1	.09	$1 \times .09 = .09$
54	2	.08	$2 \times .08 = .16$
55	3	.07	$3 \times .07 = .21$
56	4	.05	$4 \times .05 = .20$
57	5	.03	$5 \times .03 = .15$
58	6	.01	$6 \times .02 = \underline{.06}$
			$E(N) = .87$

Finally the sum of the carrying and backorder costs is entered in row 6.

$$4 + 21.75 = \$25.75$$

This series of computations is performed for each possible reorder level, and whichever reorder level displays the lowest cost in row 6 of Table 14-2 would be selected as the best strategy r^*. It appears that the best strategy for our example is to set the reorder level at $r^* = 56$. This implies a safety stock level of

$$s = r - \bar{u} = 56 - 50 = 6 \text{ units}$$

Furthermore the total cost for this strategy can be computed from Eq. 14-1.

$$\text{TC} = \frac{1{,}000}{200}(40) + \frac{200}{2}(10)(.20) \quad 6(10)(.20) + \frac{1{,}000}{200}(.05)(5)$$

$$\text{TC} = 200 + 200 + 12 + 1.25$$

$$\text{TC} = 413.25$$

We cannot say, however, that the total cost of this strategy is lower than the total cost of any other. Recall that in the process of solving Eq. 14-1 for Q and r we simplified matters by solving the first two terms for Q and the last two for r. By using only these last two terms to determine r, we have implied that the expected number of shortages per time period (year) is strictly a function of r. This is only part of the picture, for r will specify the expected number of backorders per *cycle*, and the expected number of shortages per year depends on the product of the expected number of shortages per cycle *and* the number of orders per year. Therefore the number of backorders that occurs per year depends on both Q and r, and in our simplified approach we have chosen Q^* independently of r^*. What is the magnitude of the error? You may recall that in the section on sensitivity analysis, we found that the total-cost function was relatively insensitive in the vicinity of its minimum. It is therefore reasonably safe to conclude that our simplified method for calculating Q^* and r^* will usually result in values which are negligibly off from the optimal ones.

FLUCTUATING LEAD TIME AND DEMAND

When the lead time as well as the demand is subject to risk, the problem becomes even more complex. For many items, however, this new source of variability is very small in relation to the average lead time. A reasonable approach, then, is to take the maximum likely delivery time instead of using the average.

AGGREGATE EFFECTS

It has been implied up to this point that the EOQ for each stockkeeping unit is determined in the absence of any constraints. Several constraints, in fact, may be present. Consider the situation where there is a limited fund of cash which can be used for inventory. To order the amount specified by the EOQ formula for each stockkeeping unit may result in the violation of the budget constraint. By utilizing lagrangian multiplier techniques, a feasible ordering policy can be determined.[1]

Lagrangian techniques are used to minimize (or maximize) nonlinear functions subject to a set of constraints. Since our total-cost expression is nonlinear, these techniques must be used rather than linear programming methods.

There are other constraints that could affect the quantity ordered. Warehouse space may be limited, and if the EOQ formula is used for each item, there may not be enough storage space when all the orders are received. Another constraint may be imposed by the order processing department. Perhaps they can process only a limited number of orders per period. If any of these constraints must be considered, lagrangian techniques should be used to determine the order quantity for each item such that in the aggregate these orders will not violate the constraints.

INVENTORY SYSTEMS

THE FIXED ORDER QUANTITY SYSTEM

There are two basic kinds of inventory systems in use. The first is the fixed order quantity system and the second is the periodic review system. The fixed order quantity system is precisely what we have been studying in this chapter. Whenever the reorder point r^* is reached, a fixed quantity Q^* is ordered. This requires the continuous reporting of inventory transactions, a practice which has become economical and commonplace with the use of large-scale computer systems.

THE PERIODIC REVIEW SYSTEM

The periodic review system is quite unlike the fixed order quantity system. Rather than using continuous reporting of inventory transactions, periodic checks of inventory levels are made at fixed intervals. The quantity on hand is compared with the desired inventory level, and the difference is ordered. Therefore, in the periodic review system, the order interval is fixed but the quantity ordered varies from one order to the next.

---

[1] Martin K. Starr and David Miller, *Inventory Control: Theory and Practice*, Prentice-Hall, Inc., Englewood Cliffs, N.J., 1962, p. 97.

In general it can be stated that the safety stocks for the periodic review systems will be larger than necessary for fixed order quantity systems. This results from the fact that the safety stock must provide protection for demand during a review period plus the next lead time.

Clerical costs, in the past, have been somewhat lower for this system, since only periodic audits need be taken. This advantage, however, has been rapidly declining as computer-based recording systems have become available. Fixed order systems will, however, quite likely continue to dominate the field.

A major benefit from this system is that all orders are triggered at the same time. Orders for many different items from one supplier can therefore be grouped. Ordering and shipping would therefore be much more efficient. One can easily see how supermarkets could benefit from this approach.

SUMMARY

Perhaps one of the most difficult aspects of inventory control is the collection of input data. These include demand forecasts, inventory carrying charges, fixed order costs, backorder costs, lost sales costs, and data provided by a continual monitoring of inventory levels. Most require a sizable expenditure of effort to obtain.

Since a forecast must be determined for each stockkeeping unit, this could be a particularly troublesome aspect of the input data problem. In the two previous chapters we explored forecasting techniques. The most popular of the techniques for inventory demand forecasting seems to be exponential smoothing. This, together with the appropriate inventory model, can represent a very effective system.

The other data, including inventory carrying charges, fixed order costs, backorder costs, and lost sales costs, are not generally available from accounting data. In some cases, however, such as inventory carrying and fixed order costs, the accounting data might be a starting point. The estimation of stockout costs, on the other hand, usually must rely on seasoned judgment.

Finally a mechanism must be established for the continual monitoring of inventory levels. Usually this takes the form of remote monitoring devices into which all inventory transactions are entered. This, in turn, is entered into the computer's permanent file.

The choice of the appropriate inventory model, as you have seen, requires some thought. Is demand known with certainty? If so, the simple EOQ model can be used. Perhaps price breaks are available. If so, the price-break model should be used. What if backorder costs are not too high? Then it would be appropriate to use the backorder model.

How about the level of safety stock for the case of fluctuating demand? We found two approaches for setting this level. The first was indirect and required the specification of a performance level. When this level was specified, the decision maker had in mind a stockout cost but was somehow unable to estimate it explicitly. The second approach was to estimate the stockout cost and use it directly in setting the reorder level. This model assumed that the stockout resulted in a backorder rather than in a lost sale. Models with other assumptions can, of course, be built. Once again it should be clear that the understanding of the assumptions is of paramount importance before the model can be used.

QUESTIONS

Q14-1. Distinguish between certain demand, uncertain demand, and demand under risk.

Q14-2. What difference is there between inside and outside procurement in their effect on production planning and inventory control?

Q14-3. What is the difference in consequence between an overstock in a dynamic and in a static inventory problem?

Q14-4. Why is it unnecessary to consider total procurement costs in the total-cost expression of the simple EOQ model?

Q14-5. Differentiate between the fixed order quantity and periodic review systems.

Q14-6. When demand is subject to risk, and backorder costs are known, what is the method for setting the reorder point? Why is it an approximation?

Q14-7. What is the consequence of setting the safety stock too high or too low?

Q14-8. Why is the cost of a stockout difficult to compute?

Q14-9. Is the simple EOQ model sensitive to forecast errors? Explain.

Q14-10. If an electronics manufacturer felt that possible advances in the state of the art precluded the carrying of large inventories, how could he incorporate this in his inventory-ordering strategy?

PROBLEMS

P14-1. The yearly demand for stock item 17096 is estimated to be 4,000 units. Fixed order costs are \$40 per order, while per-piece costs are \$10. If inventory carrying costs are 10 percent per unit per year, compute the following:

(*a*) Economic order quantity
(*b*) Total cost of the EOQ strategy

P14-2. If demand was estimated incorrectly as 4,000 in Prob. 14-1 but actually turned out to be 5,000, what should the EOQ have been? What was the percentage error in total costs?

P14-3. The following estimates have been determined for an inventory problem.

$D = 2{,}000$ units
$A = \$15$ per order
$I = 15\%$
$C = \$10$ per unit

If the unknown but actual fixed order and inventory carrying costs are \$20 per order and 10 percent per period, what will be the magnitude of the order quantity and the total-cost errors?

//// **P14-4.** If carrying costs have been estimated to be 120 percent of the unknown actual, and fixed order costs have been estimated to be 80 percent of the actual, what will be the percentage error in total costs? ////

4-5. Consider the following price schedule:
If the forecasted demand is 8,000 units, the fixed order costs \$100, and the inventory carrying costs 25 percent, what quantity should be ordered?

If $Q < 850$, then \$10 per unit
If $Q \geq 850$, then \$8 per unit

P14-6. Consider the following price schedule:

If $Q < 700$, then \$10 per unit
If $Q \geq 700$, then \$7 per unit

If the forecasted demand is 8,000 units, the fixed order costs \$100, and the inventory carrying costs 25 percent, what quantity should be ordered?

P14-7. Given the following information, determine if the price break should or should not be taken.

A = \$10 per order
I = 10% per unit per year
D = 500 units per year

If $Q < 250$, C = \$10 per unit
If $Q \geq 250$, C = \$2.50 per unit

P14-8. The forecasted demand for a particular stock unit is 40,000 units. The fixed order costs are \$25 per order and the inventory carrying costs are 25 percent per unit per year. The manufacturer offers price breaks at the following two points:

If $0 \leq Q < 100$, then C = \$8
If $100 \leq Q < 800$, then C = \$7
If $Q \geq 800$, then C = \$6

What quantity should be ordered?

P14-9. Draw a flow chart for the decision process when two price breaks are offered.

P14-10. The following estimates have been made for an inventory problem:

D = 2,000 units
A = \$15 per order
I = 15% per period

Given the following price schedule, how many units should be ordered?

If $Q < 300$, then C = \$10
If $300 \leq Q < 500$, then C = \$9
If $Q \geq 500$, then C = \$8

P14-11. The cost for incurring a backorder is estimated to be \$10 per backorder. If demand is estimated to be 4,000 units, fixed order costs \$10 per order, inventory carrying cost 10 percent per unit per year, and per-piece cost \$1, find the economic order quantity and the maximum level of inventory.

P14-12. Solve the following problem for order quantity:

D = 3,000 units
A = \$10 per order
I = 10% per unit per year
C = \$2 per unit
B = \$5 per backorder

P14-13. The demand during several lead time periods has been observed and recorded below.

Lead time demand	*Relative frequency*
110	.05
111	.10
112	.15
113	.25
114	.20
115	.10
116	.08
117	.07

At what point should the reorder level be set so that the chances of a stockout are 15 percent or less?

P14-14. The lead time demand distribution for a particular product can be approximated by a normal distribution with a mean of 250 and a standard deviation of 5. At what level should the reorder point be set so that there will be less than a 7 percent chance of a stockout? What will the safety stock be at this level?

P14-15. The probability distribution for lead time demand is estimated to be the following.

Lead time demand	*Probability*
30	.07
31	.15
32	.26
33	.18
34	.12
35	.08
36	.07
37	.07

In addition, demand is 2,000 units, fixed order costs $25 per order, per-piece cost $2 per unit, inventory carrying costs 25 percent per unit per year, and backorder costs $1 per unit backordered.

What is the optimal reorder level, the size of the safety stock, and the total cost of this strategy?

✓ **P14-16.** What is the optimal reorder level, the size of the safety stock, and the total cost of the optimal strategy for the following inventory problem?

Lead time demand	*Probability*
48	.02
49	.03
50	.06
51	.07
52	.20
53	.24
54	.20
55	.07
56	.06
57	.03
58	.02

D = 1,800 units
A = \$30 per order
I = 15% per unit per period
B = \$1 per unit backordered
C = \$2 per unit

P14-17. Below are given the probability distributions of *daily* demand, lead time, and backorders. The backorder table represents the likelihood that a stockout will become a lost sale. For example, if a stockout has existed for 3 days, the chances it will become a lost sale are 80 percent.

Probability distribution of daily demand:

Number of units demanded	*Probability*
7	.10
8	.20
9	.25
10	.20
11	.15
12	.10

Lead time distribution:

Lead time (days)	*Probability of occurrence*
4	.20
5	.40
6	.30
7	.10

Probability of a stockout turning into a lost sale:

Number of periods backordered	*Probability*
0	.20
1	.40
2	.65
3	.90
4	1.00

In addition, the carrying costs per day are estimated to be \$.10 per unit, the cost of a lost sale is \$30 per unit while the cost of a backorder is \$5 per unit, fixed order costs are \$30 per order, and per-piece costs are \$1 per unit.

How would you simulate a strategy with a reorder point of 60 units and an order quantity of 300? Is this necessarily the best strategy? What about other strategies?

BIBLIOGRAPHY

Buffa, Elwood S.: *Production-Inventory Systems: Planning and Control*, Richard D. Irwin, Inc., Homewood, Ill., 1968, chaps. 3 and 4.

Magee, John F., and David M. Boodman: *Production Planning and Inventory Control*, McGraw-Hill Book Company, New York, 1967.

Niland, Powell: *Production Planning, Scheduling, and Inventory Control*, The Macmillan Company, New York, 1970.

Starr, Martin K.: *Systems Management of Operations*, Prentice-Hall, Inc., Englewood Cliffs, N.J., 1971, chap. 12.

——— and David W. Miller: *Inventory Control: Theory and Practice*, Prentice-Hall, Inc., Englewood Cliffs, N.J., 1962.

Wagner, Harvey M.: *Principles of Operations Research*, Prentice-Hall, Inc., Englewood Cliffs, N.J., 1969, chaps. 9 and 19 and Appendix II.

Whitin, T. M., and G. Hadley: *Analysis of Inventory Systems*, Prentice-Hall, Inc., Englewood Cliffs, N.J., 1963.

15

Loading, Sequencing, and Control

INTRODUCTION

When the aggregate planning problem was discussed in Chapter 13, we did not work our way down to the planning of production for individual items. The concern at that stage was to establish an aggregate strategy within which these detailed schedules could eventually be established.

After the mix of jobs which will flow through the process is finalized, it is necessary to compile a schedule specifying when jobs will be started, the machines on which they will be processed, the order in which they will be sequenced on these machines, and their expected completion date. This task is usually initiated one month or less before processing is begun.

The two major problems at this stage are loading and sequencing. Loading is concerned with the assignment of jobs to work centers, while sequencing focuses on the order in which these jobs are processed. First we look at the loading problem.

LOADING

The focus of attention in loading is the total load placed on a particular work center. The traditional approach for handling this problem has been to assign work to a center and to monitor the load with a Gantt load chart. An example of this chart is shown in Fig. 15-1. Here we can see the cumulative work load that has been placed on the center as well as the expected week-by-week load.

Although the Gantt load chart is a useful record-keeping device, it is not particularly helpful when there exist alternative machines on which the job can be loaded. Under these circumstances, a fairly complex decision problem exists. Let us take a closer look.

Consider a simple case where five jobs have to be assigned to five machines. Each job must be assigned to a different machine because these jobs are expected to run continuously. Furthermore, each job can be performed on any of the five machines. The cost of assigning job i to machine j is shown in Fig. 15-2. The problem is which job to assign to what machine. Since the rim conditions are equal to 1, it can be solved as an assignment problem.

The point that this simple example should bring out is that when there is more than one way in which to load a job, a loading strategy must be found

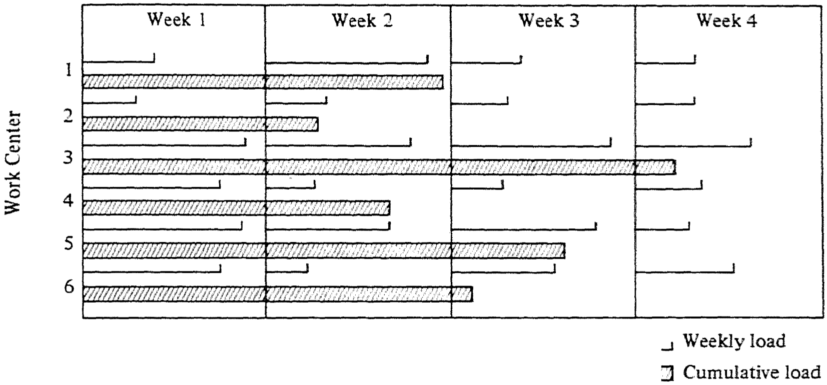


Fig. 15-1 Gantt chart.

	Machine j				
	1	2	3	4	5
	40	27	19	36	27
	32	39	41	26	17
ob 3	16	17	27	20	31
	51	19	39	14	16
	43	21	16	29	35

Fig. 15-2 Cost of assigning job i to machine j.

which minimizes the cost of processing for all jobs under consideration. In the Gantt load chart approach, the choice of where to load a job, if there is a choice, is left to the discretion of the scheduler. Therefore, it is more of a record-keeping device than a decision-making tool. When used alone, it skirts the main problem.

In a more frequent and more complex case, a set of jobs must be assigned to machines; however, more than one job can be assigned to each machine. Sometimes it is even acceptable for the job to be split between two or more machines. For example, a job requiring the production of 1,000 pieces may have 300 of them run on one machine and 700 run on another. The constraints imposed on this loading problem include machine and labor capacity. The objective may be stated as either the minimization of process costs, the maximization of equipment utilization, the minimization of job lateness, or one of several others. The full development of this problem, however, is beyond the scope of this chapter, and the interested reader is referred elsewhere.[1]

For the remainder of this chapter we shall assume that loading is a fairly straightforward problem. That is, either there is but one loading alternative for each operation, or an acceptable loading solution already exists. This will allow us to focus on the sequencing problem.

SEQUENCING

Once the load on a machine center has been determined, the next problem is to determine the sequence in which these jobs should be executed.

Although this problem was originally studied in a manufacturing environment, the sequencing problem is a more general phenomenon. In a hospital, a group of patients must be sequenced through a number of tests; at the airport, a number of aircraft holding in a landing pattern must be sequenced for landing; at a bank, the customers must be sequenced through a loan department; at a service station, a group of automobiles must be sequenced through a series of work centers; and a salesman must determine in what sequence he will call on 10 customers.

In almost all of these situations, the problem is solved without our even realizing that a problem exists. Consider the problem faced by the salesman. He must decide the order in which to call on his customers, who are located in 10 cities. There are many possible sequences which he could follow. If we numbered the cities 1 through 10, the following represent two possible sequences:

1-2-3-4-5-6-7-8-9-10

1-3-2-4-5-8-7-10-6-9

In fact there are 10! or over 3 million possible sequences. Suppose that the salesman wanted to minimize driving time. Then for each of the 3 million sequences, a driving time could be computed. Finally, the one with the lowest time would be chosen as the sequence to be followed.

---

[1] For a good introduction to the subject of loading see Martin K. Starr, *Systems Management of Operations*, Prentice-Hall, Inc., Englewood Cliffs, N.J., 1971, chap. 10.

It can be seen from this example that a problem which is usually seen as trivial from the salesman's point of view turns out to be quite substantial from the management scientist's point of view. To completely enumerate all possible sequences and from this set choose the best one is completely untenable, even for this small problem.

In this chapter we will focus on more efficient ways of solving this and other sequencing problems. First we will develop a framework into which this analysis will be placed. Then we will develop a series of sequencing models, each to be used under different conditions. Finally two case studies are presented which illustrate approaches to the sequencing and control of complex systems.

A FRAMEWORK FOR SEQUENCING PROBLEMS

The models that will be presented will differ in their consideration of one or another of the following factors:

1. The number of jobs n
2. The number of machines m
3. The type of facility: flow shop—job shop
4. Manner in which jobs arrive at the facility: static or dynamic
5. Criterion by which alternatives will be evaluated

The first factor is the number of jobs, n, that must be processed through the facility, and the second is the number of machines that will be utilized. The third factor concerns the order in which machine centers appear in the routing of individual jobs. If all jobs follow the same path, that is, if they go through the same sequence of machines, we have a continuous or *flow shop*. If, at the other extreme, there is no common pattern for the flow of jobs through the shop, a *job shop* pattern is said to exist. The fourth factor is the manner in which jobs arrive at the facility. If the jobs arrive simultaneously, the shop is idle upon receipt of these orders, and no further jobs arrive, a *static arrival pattern* exists. This kind of situation requires that we focus only on a set of fixed jobs. A more difficult situation, but one which is usually more realistic, is one where jobs are continuously arriving at a facility. This is the case of a *dynamic arrival pattern*. Finally we turn to the criterion by which alternative sequencing plans are compared. Here there is certainly no agreement as to the best one. These include the utilization of equipment as measured by the processing time, the average lateness of orders, the total setup time, and the minimization of work-in-process, to name just a few.

n JOBS, ONE MACHINE, FLOW SHOP, STATIC ARRIVAL PATTERN, SPT

The first case which will be examined assumes that any number of jobs n are queued up at *one* machine. Since all the jobs must go through this one machine, we essentially have a flow shop problem. In addition we assume that all of the jobs are received at once and no new jobs arrive. Therefore a static arrival pattern prevails. The objective will be to find the shortest average processing time (SPT) for these jobs, where processing time is the sum of waiting and operation time. If we let C_i be the time in which each job is completed, for n jobs the average processing time is:

$$\bar{C} = \frac{\sum_{i=1}^{n} C_i}{N}$$

Since N is constant for all sequences, the sequence which minimizes the sum of the processing times will be identical to that which minimizes the average processing time.

The operation times for six jobs are given in Table 15-1. The problem is to find that sequence which minimizes average processing time $\bar{C}$.

Consider the trial solution A-B-C-D-E-F. Its average processing time is:

$$\bar{C} = \frac{7 + 13 + 17 + 20 + 22 + 23}{6} = 17 \text{ min}$$

Another trial solution is sequence F-E-D-C-B-A. Its average processing time is:

$$\bar{C} = \frac{1 + 3 + 6 + 10 + 16 + 23}{6} = 10 \text{ min}$$

Clearly the choice of a sequence has an influence on average processing time per job.

In Fig. 15-3 a graph is drawn for the trial solution A-B-C-D-E-F. Along the abscissa is time in minutes and along the ordinate, at one-unit intervals, is the progression from N to 1, within which the jobs will be placed.

In this trial solution, the first job on the schedule is A. Therefore the rectangle that appears first in the sequence is 1 unit high by $P_A = 7$ minutes long. The second job is B. Its rectangle is 1 unit by $P_B = 6$ minutes. In similar fashion the rest of the rectangles are drawn.

Table 15-1 Job operation times

Job n	*Operation time* P_n
A	7 min
B	6
C	4
D	3
E	2
F	1

The total area of this graph, including both the shaded area and the rectangles, represents the sum of the job processing times. Now if we were to generate another trial sequence, the sum of the area *within the rectangles* would not change. That is, this area is independent of the sequence. The *shaded area*, however, is sequence-dependent. Consider, for example, the sequence for the second trial solution: F-E-D-C-B-A. Its graph is shown in Fig. 15-4. Clearly the shaded area for this sequence is much less than the shaded area in the preceding sequence.

THE MINIMIZATION OF AVERAGE PROCESSING TIME

We can characterize each block by an arrow on its diagonal. Each schedule graph will therefore be a head-to-tail chain of these arrows whose slope is $-1/P_n$. It therefore seems reasonable to say that the shaded area is minimized when the arrows form a convex curve. This is equivalent to sequencing the jobs by nondecreasing operation times, or: $P_{[1]} \leq P_{[2]} \leq P_{[3]} \leq P_{[4]} \leq \cdot \cdot \cdot \leq P_{[n]}$.

If the jobs in this example were arranged by nondecreasing operation times, we would have $P_F \leq P_E \leq P_D \leq P_C \leq P_B \leq P_A$, and they would be scheduled in the following order: F-E-D-C-B-A. For this sequence the arrows form a convex

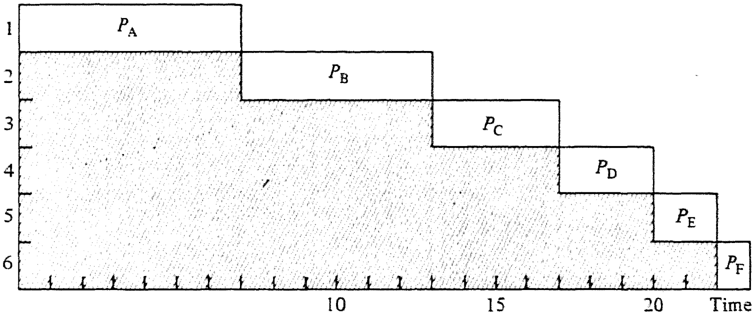


Fig. 15-3 Trial solution A-B-C-D-E-F.

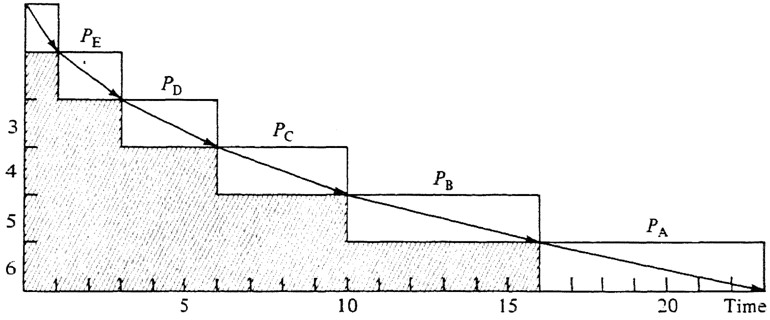


Fig. 15-4 Schedule graph of optimal solution F-E-D-C-B-A.

curve, the shaded area is minimized, and we therefore have a schedule which minimizes the average processing time.

As an illustration of how this rule could be applied, consider the following familiar situation. In a supermarket a group of n people (n jobs) is waiting in line at a checkout counter (1 machine). The sequence in which they will be processed through the system is usually based on when they arrived: first come, first served. That this rule is not the best should be clear. If the first customer in line has a large order, this will increase the average processing time per customer. We just learned that to minimize average processing time the customers (or jobs) should be sequenced by nondecreasing operation time. This means that the customer with the fewest items goes first, then the customer with the next fewest, and so on. Although the supermarkets have never tried this, they have established checkout lanes for fewer than 10 or so items. This will, of course, decrease the average processing time somewhat.

OTHER CRITERIA

It seems that one of the most important criteria by which practitioners schedule jobs in a typical job shop is the minimization of job lateness. By employing the SPT rule, surprisingly enough the *average* job lateness is also minimized. The reason why this is surprising is that due dates were not considered by this rule in any way.

If, however, the criterion is to minimize the *maximum* lateness of any job in the set of n jobs, it can be shown that the optional sequence is in order of nondecreasing due dates: $d_{[1]} \leq d_{[2]} \leq d_{[3]} \leq \cdot \cdot \cdot \leq d_{[n]}$. Proof of this can be found elsewhere.[2] The point that should be made is that the particular rule which should be used depends upon the criterion chosen. Rarely will one rule satisfy all relevant criteria.

n JOBS, ONE MACHINE, FLOW SHOP, STATIC ARRIVAL PATTERN, SEQUENCE-DEPENDENT SETUP TIMES

In the previous model an unwritten assumption was that if a setup was necessary for each job, the time which it took was independent of the previous job. This being the case, the setup times can simply be included in the operation time P_n. For example, in the problem illustrated in Table 15-1, the operation time for job A might comprise 1 hour of setup time and 6 hours of run time. By expressing the operation time in this way it is implied that the setup time is independent of the job by which it is preceded. For job A the setup and run time is 7 hours regardless of whether it was preceded by job B, C, D, E, or F. There are several occasions, however, when the setup time depends on the previous job. Consider a scheduling problem found in a large printing shop. The length of time required to set up a job depends upon the color of the ink and the paper size of the previous job. The actual setup and run times for a batch of six jobs which must be processed each month are shown in Fig. 15-5. For example, if job 1 is followed by job 5, the setup and run time for job 5 will be 14 hours; if job 5 is followed by job 2, the setup and run time will be 7 hours; and so on. Notice that the matrix is not symmetrical. The setup and run time required to go from job 1 to job 5 is not the same as to go from job 5 to job 1. The problem is to find a sequence for these six jobs which minimizes the total setup and run time. There are several other situations in which a similar sequencing problem can be found. In the chemical and processing industries different products are often processed through the same sequence of operations. After each one is processed, the equipment must be cleaned and set up for the next job. The cleaning and setup times are frequently dependent on both the product which has just been processed and the next product. Other examples include the manufacture of ice cream, paint, petroleum products, and industrial chemicals.

---

[2] Richard W. Conway, W. Maxwell, and L. Miller, *Theory of Scheduling,* Addison-Wesley Publishing Company, Reading, Mass., 1967, p. 30.

Another example which exhibits this form is the "traveling salesman problem." In this problem a salesman must visit each of n cities once and only once and finally return to his point of origin. In choosing the best sequence, either cost, time, or distance may be minimized. It is seldom that any one schedule will minimize all three.

THE DIFFERENCE BETWEEN THIS AND THE ASSIGNMENT PROBLEM

In some respects these problems seem to resemble the assignment problem. A closer look, however, will show us that this is not actually the case. Consider the problem presented in Fig. 15-5. If we let

$$X_{ij} = \begin{array}{ll} \text{if } i \text{ to } j \text{ is an included sequence} & i = 1, 2, \ldots, 6 \\ \text{otherwise} & j = 1, 2, \ldots, 6 \end{array}$$

and c_{ij} = cost of the sequence ij

To job

		!	3	4	5	
	–	1	7	3	14	2
	3	–	6	9	1	24
	6	14	–	3	7	3
From 4	2	3	5	–	9	11
	15	7	11	2	–	4
	20	5	13	4	18	–

Fig. 15-5 Job setup and run times. (*From Conway et al., Theory of Scheduling, Addison-Wesley Publishing Company, Reading, Mass., 1967, p. 58.*)

the objective function can be written as:

$$\text{Minimize} \sum_{i=1}^{6} \sum_{j=1}^{6} c_{ij}X_{ij}$$

The constraints will include:

$$X_{ij} = 1 \qquad j = 1, 2, \ldots, 6$$

and

$$\sum X_{ij} = 1 \qquad i = 1, 2, \ldots, 6$$

Up to this point the mathematical specification of the problem is exactly that of the assignment problem. There is, however, one important restraint that must be added. *No tour (schedule) can return to its starting place until all n cities (jobs) have been visited (assigned).* For example, the matrix shown in Fig. 15-6 is a feasible solution to an assignment problem; but if we look at it closely, it is an infeasible solution to the sequencing problem presented in Fig. 15-5. Although the conventional rim conditions associated with the assignment problem are not violated, the additional restraint which we imposed is violated. Namely, the solution includes the following cyclic subtours:

1-6-1 and 2-5-2 and 3-4-3

Consequently, this problem is more complex to solve than the assignment problem. We now turn to a computational procedure for solving it. You will find it quite similar but not identical to the assignment algorithm presented in Chapter 8. Essentially the steps involve the partitioning of the problem into less complex problems, the establishment of lower bounds, and the check to ensure that no cyclic subtours are included. The process continues with additional partitioning, and so on, until a solution is reached.

					1
				1	
			1		
		1			
	1				
1					

Fig. 15-6 Trial solution matrix.

REDUCING THE MATRIX BY ROWS AND COLUMNS

The first step in the search for the optimal solution is to reduce the matrix *first by rows* and *then by columns*.[3] The reductions that can be made to each row are shown in Fig. 15-7 while the column reductions are shown in Fig. 15-8. The new matrix, which has at least one zero in every row and column, is shown in Fig. 15-9. Since we have reduced the original matrix by $13 + 3 = 16$, this will establish a lower bound on the problem. That is, any solution to the problem will have a cost of at least 16. If the zero cells are so placed that they represent a feasible sequence for the six jobs, the optimal solution has been reached and its cost would be 16. If they are not, nonzero

	–	1	7	3	14	2	1
	3	–	6	9	1	24	1
3	6	14	–	3	7	3	3
4	2	3	5	–	9	11	2
5	15	7	11	2	–	4	2
6	20	5	13	4	18	–	4
							13

Fig. 15-7 Row reduction.

1	–	0	6	2	13	1
2	2	–	5	8	0	23
3	3	11	–	0	4	0
4	0	1	3	–	7	9
5	13	5	9	0	–	2
6	16	1	9	0	14	–
	0	0	3	0	0	0

Fig. 15-8 Column reduction.

–	2 0	3	2	13	1
2	–	2	8	6 0	23
3	11	–	0	4	1 0
2 0	1	2 0	–	7	9
13	5	6	2 0	–	2
16	1	6	1 0	14	–

Fig. 15-9 Reduced matrix.

---

[3] The problem can also be solved by reducing first by columns and then by rows.

cells must be used, increasing the minimum-cost solution from this lower bound. The zero cells in Fig. 15-9 are not located in such a way that a feasible schedule can be identified by using only those cells. Consequently we must employ a branch-and-bound algorithm for solving this problem. We now turn to this algorithm and the task of making the first assignment.

FIRST ASSIGNMENT

Above each zero entry is entered the opportunity cost of not making that particular assignment. For example, if an assignment is not made in cell 1-2, cell 1 must be terminated somewhere else. In the interest of cost minimization, that assignment would be made to cell 1-6. In addition, cell 2 must originate somewhere else besides cell 1. It would be assigned to cell 4-2. Therefore, if cell 1-2 is not assigned, a minimum additional cost of 1 + 1 = 2 will be incurred. In a similar fashion the remaining zero cells are evaluated.

Next an assignment is made to the cell with the highest opportunity cost. As you will recall from Chapter 8, the assignment is made here to avoid incurring this high cost. Accordingly, we make our first assignment to cell 2-5, and it serves as the basis for the partition of the original problem into two less complex ones. First we partition into all those problems which have cell 2-5 as an included sequence; then we partition into all those problems which do not allow 2-5 as an included sequence.

THE FIRST TWO BRANCHES

In Fig. 15-10 the first new problem is generated in which cell 2-5 is assigned. An infinite cost, represented by dashes, is entered in the remaining cells of row 2, since job 2 will be followed *only* by job 5. Dashes are also entered in the remaining cells of column 5, since job 5 can be preceded only by job 2. We also enter a dash in cell 5-2, since 5-2-5 would result in a prohibited cyclic subtour. Now the remaining costs are entered into the matrix.

The distinguishing feature of the second new problem is that the assignment

—	0	3	2	—	1
—	—	—	—	X	—
3	11	—	0	—	0
0	1	0	—	—	9
13	—	6	0	—	2
16	1	6	0	—	—

Fig. 15-10 First assignment, first new problem: S_{25}-(16).

to cell 2-5 is prohibited. To accomplish this a dash is entered in its cell. The new problem is presented in Fig. 15-11.

Both of these new problems should be examined for possible row and column reductions. The first new problem cannot be reduced any further; it therefore has a lower bound of 16, the same as the original problem. This new problem will be denoted by S_{25} — (16), where the subscript represents the assigned cells and the parenthesis represents the lowest bound.

The second new problem can be reduced by 2 in the second row. The new matrix after this reduction is shown in Fig. 15-12. Checking for possible column reduction, we see that column 5 can be reduced by 4. The new fully reduced matrix is shown in Fig. 15-13. The lower bound for this problem, however, has increased from 16 to 22, since row and column reductions of 2 + 4 = 6

–	0	3	2	13	1
2	–	2	8	–	23
3	11	–	0	4	0
0	1	0	–	7	9
13	5	6	0	–	2
16	1	6	0	14	–

Fig. 15-11 Second new problem.

–	0	3	2	13	1
0	–	0	6	–	21
3	1	–	0	4	0
0	1	0	–	7	9
13	5	6	0	–	2
16	1	6	0	14	–

Fig. 15-12 Second new problem after row reduction.

–	0	3	2	9	1
0	–	0	6	–	21
3	1	–	0	0	0
0	1	0	–	3	9
13	5	6	0	–	2
16	1	6	0	10	–

Fig. 15-13 Second new problem fully reduced: S_{n25}-(22).

were undertaken. We can therefore denote this problem by S_{n25} — (22), where the subscript $n25$ denotes that cell 2-5 is *not* assigned.

Let us take a more general view of what we have just accomplished. We started with a matrix which we can say represented all possible solutions to the problem. It had a lower bound of 16. From this solution set we partitioned into all those problems that included 2-5 and those problems that did not include 2-5 as an assigned set. This can be viewed graphically, as shown in Fig. 15-14.

We will now branch out from our new problems, each time reducing the size of the problem until it is eventually solved. The decision now, however, is from which of these two new problems to branch. First let us try S_{25} — (16), since its lower bound is less and it has more assignments than does S_{n25} — (22).

BRANCHING FROM S_{25} — (16)

Figure 15-10 is the matrix for the set of problems S_{25} — (16). It is repeated in Fig. 15-15, and in addition we enter the opportunity costs for the zero cells. Using the maximum opportunity cost as a basis for the partition, we choose cell 4-1. First we partition into all problems including cell 2-5 *and* cell 4-1. This is accomplished in Fig. 15-16. We also assign an infinite cost to assignments other than cell 4-1 in column 1, assignments other than cell 4-1 in row 4, and assignment of cell 1-4. Finally the remaining cell values are entered. It is not possible to reduce the rows of the matrix; however, the third column can be reduced by 3. Therefore, the lower bound of this new

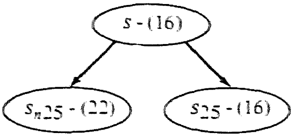


Fig. 15-14

1	–	2 0	3	2	–	1
2	–	–	–	–	☒	–
3	3	11	–	0 0	–	1 0
4	3 0	1	3 0	–	–	9
5	13	–	6	2 0	–	2
6	16	1	6	· 1 0	–	–

Fig. 15-15 Branching from S_{25}-(16).

1	2	3	4	5	6
–	0	3	–	–	1
–	–	–	–	☒	–
–	11	–	0	–	0
☒	–	–	–	–	–
–	–	6	0	–	2
–	1	6	0	–	–

Fig. 15-16

problem is $16 + 3 = 19$, and we can denote it as $S_{25,41} - (19)$. The reduced matrix is shown in Fig. 15-17.

Now we must partition into those problems which do *not* include cell 4-1. This is accomplished by entering a dash in cell 4-1 of Fig. 15-18. Next we reduce the rows and columns. Only column 1 can be reduced by 3. The reduced matrix is shown in Fig. 15-19. It has a lower bound of $16 + 3 = 19$ and can be denoted by $S_{25,n41} - (19)$. The sequence of partitions up to this point is shown in Fig. 15-20.

BRANCHING FROM $S_{25,41} - (19)$

Branching can now take place from one of three places: $S_{n25} - (22)$, $S_{25,n41} - (19)$, or $S_{25,41} - (19)$. The most promising is $S_{25,41} - (19)$, since it has a lower bound no greater than the other two and it has more assignments specified than $S_{25,n41} - (19)$. Although we will not branch from these other two matrices at

–	0	0	–	–	1
–	–	–	–	☒	–
–	11	–	0	–	0
☒	–	–	–	–	–
–	–	3	0	–	2
–	1	3	0	–	–

Fig. 15-17 $S_{25,41}$-(19).

1	2	3	4	5	6
–	0	3	2	–	1
–	–	–	–	☒	–
3	11	–	0	–	0
–	1	0	–	–	9
13	–	6	0	–	2
16	1	6	0	–	–

Fig. 15-18

			3	4	5	6
1	–	0	3	2	–	1
2	–	–	–	–	X	–
3	0	11	–	0	–	0
4	–	1	0	–	–	9
5	10	–	6	0	–	2
6	13	1	6	0	–	–

Fig. 15-19 $S_{25,n41}$-(19).

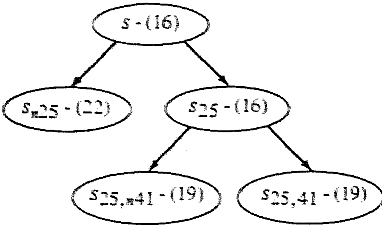


Fig. 15-20

			3	4	5	6
1	–	1 0	3 0	–	–	1
2	–	–	–	–	X	–
3	–	11	–	0 0	–	1 0
4	X	–	–	–	–	–
5	–	–	3	2 0	–	2
6	–	1	3	1 0	–	–

Fig. 15-21 Branching from $S_{25,41}$-(19).

this time, we will return to them later to ensure that they do not lead to better solutions than the one toward which we are now heading.

To partition from $S_{25,41}$ — (19), we return to Fig. 15-17 and enter the opportunity costs in the zero cells. The results are shown in Fig. 15-21. Partitioning is based on cell 1-3. First we partition into those problems which include cell 1-3. In addition to the usual inadmissible cells, a dash must be placed in cell 3-4, since its inclusion would result in a cyclic subtour of less than six jobs (Fig. 15-22).

Each time an assignment is made and after the usual cells are filled with a dash, the remaining cells (those which have finite values) should be checked to see if their future selection would cause a cyclic subtour of less than six jobs. For example, in this case if cells 3-2, 3-4, 3-6, 5-4, 5-6, 6-2, and 6-4 are checked, you will find that cell 3-4 when added to the partially determined tour (2-5 and 4-1) will result in the subtour 1-3-4-1. Since this must be avoided, a dash is entered in cell 3-4. Essentially it is this checking procedure which differentiates this algorithm from the assignment algorithm presented in Chapter 8.

After the relevant costs and dashes are entered into Fig. 15-23, the matrix is reduced. The results are shown in Fig. 15-24. Its lower bound is $19 + 1 = 20$ and it can be denoted by $S_{25,41,13}$ — (20).

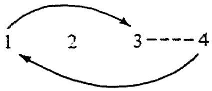


Fig. 15-22

–	–	☒	–	–	–
–	–	–	–	☒	–
–	11	–	–	–	0
☒	–	–	–	–	–
–	–	–	0	–	2
–	1	–	0	–	–

Fig. 15-23

–	–	☒	–	–	–
–	–	–	–	☒	–
–	10	–	–	–	12 0
☒	–	–	–	–	–
–	–	–	2 0	–	2
–	10 0	–	0 0	–	–

Fig. 15-24 $S_{25,41,13}$-(20).

–	0	–	–	–	1
–	–	–	–	╳	–
–	11	–	0	–	0
╳	–	–	–	–	–
–	–	3	0	–	2
–	1	3	0	–	–
		3			

Fig. 15-25

1	2	3	4	5	6
–	0	–	–	–	1
–	–	–	–	╳	–
–	11	–	0	–	0
╳	–	–	–	–	–
–	–	0	0	–	2
–	1	0	0	–	–

Fig. 15-26 $S_{25,41,n13}$-(22).

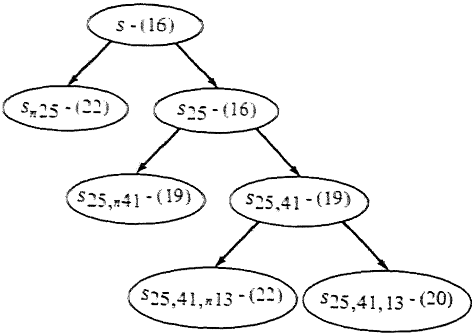


Fig. 15-27

Now the partition into problems which do not include cell 1-3 must be considered. In Fig. 15-25 the new matrix is drawn. After reduction by 3 in column 3, we have the matrix shown in Fig. 15-26. Its lowest bound is $19 + 3 = 22$, and it is denoted by $S_{25,41,n13}$ — (22). Our branching up to this point is shown in Fig. 15-27.

BRANCHING FROM $S_{25,41,13}$ — (20)

To branch from $S_{25,41,13}$ — (20), we compute the opportunity costs for the zero cells. These costs are entered in Fig. 15-24 to conserve on space. We see that partitioning will be based on cell 3-6. The new matrix is drawn in Fig. 15-28. Notice that a dash must also be placed in cell 6-4, since its inclusion in a tour would result in the following unacceptable cyclic subtour: 6 → 4 → 1 → 3 → 6. The resulting matrix cannot be reduced; therefore its lowest bound is 20 and it can be denoted by $S_{25,41,13,36}$ — (20).

The other branch, prohibiting cell 3-6, is depicted in Fig. 15-29. It can be reduced by 10 in row 3 and then by 2 in column 6. The reduced matrix is presented in Fig. 15-30. Its lowest bound is $20 + 12 = 32$, and it can be denoted by $S_{25,41,13,n36}$ — (32). Graphically, we can depict the branching up to this point by the tree shown in Fig. 15-31.

	1	2	3	4	5	6
1	–	–	✕	–	–	–
2	–	–	–	–	✕	–
3	–	–	–	–	–	✕
4	✕	–	–	–	-	–
5	–	–	–	0	–	–
6	–	0	–	–	–	–

Fig. 15-28 $S_{25,41,13,36}$-(20).

–	–	✕	–	–	–
–	–	–	–	✕	–
–	10	–	–	–	–
✕	–	–	–	–	–
–	–	–	0	–	2
–	0	–	0	–	–

Fig. 15-29

	1	2	3	4	5	6
	–	–	✕	–	–	–
	–	–	–	–	✕	–
	–	0	–	–	–	–
	✕	–	–	–	–	–
	–	–	–	0	–	0
	–	0	–	0	–	–

Fig. 15-30 $S_{25,41,13,n36}$-(32).

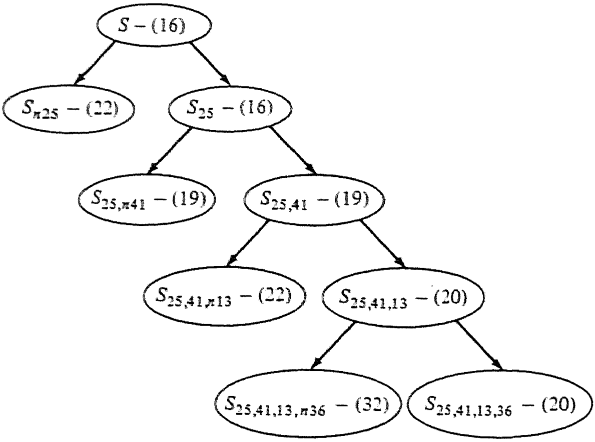


Fig. 15-31

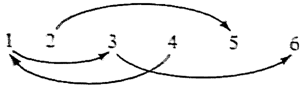


Fig. 15-32

A TENTATIVE SOLUTION

Returning to $S_{25,41,13,36} - (20)$, it can be shown that we have a solvable problem. The assignments already specified include those shown in Fig. 15-32. From this can be observed two disconnected sequences:

4-1-3-6 and 2-5

There is only one way in which they can be combined in a valid sequence:

4 1 3 6 2 5

CHECKING OTHER BRANCHES

The cost of this solution is 20. Since we now have a complete solution, we can look at some of the uncompleted branches of Fig. 15-31 to see if they are worth pursuing. For example, the branch that would emanate from $S_{n25} - (22)$ is no longer worth considering, since a solution pursued in this direction cannot have a cost lower than 22. Our present solution at a cost of 20 is preferred. We can also strike out the branches $S_{25,41,n13} - (22)$ and $S_{25,41,13,n36} - (32)$, since they both have lower bounds greater than our present solution.

There is, however, one branch which we must pursue before we can declare our present solution the optimal one. This is $S_{25,n41} - (19)$ with a lower bound of 19.

Returning to Fig. 15-19, we must proceed by computing the opportunity costs. The results are shown in Fig. 15-33. Partitioning is based on 3-1. The matrix for the branch which assigns cell 3-1 is shown in Fig. 15-34. Column 6 can be reduced by 1. This results in the reduced matrix of Fig. 15-35, with a lower bound of 20. This approach can be discarded, since it is impossible that it will lead to a solution lower than the present one of 20. At best its solution can be equal to the one we have.

Now we turn to the other branch which prohibits the assignment of cell 3-1. The matrix and its opportunity costs are shown in Fig 15-36. It can be reduced by 10 in the first column. The result is shown in Fig. 15-37. The lower bound is 19 + 10 = 29, and it is clear that this branch can also be discarded. We have, therefore, eliminated all but $S_{25,41,13,36}$ — (20), with its implied

	1	2	3	4	5	6
	–	0 ╲ 2	3	2	–	1
	–	–	–	–	╳	–
3	0 ╲ 10	11	–	0 ╲ 0	–	0 ╲ 1
4	–	1	0 ╲ 4	–	–	9
5	10	–	6	0 ╲ 2	–	2
6	13	1	6	0 ╲ 1	–	–

Fig. 15-33

–	0	–	2	–	1
–	–	–	–	╳	–
╳	–	–	–	–	–
–	1	0	--	–	9
–	–	6	0	–	2
–	1	6	0	–	–

Fig. 15-34

	–	0	–	2	–	0
	–	–	–	–	╳	–
	╳	–	–	–	–	–
	–	1	0	–	–	8
5	–	–	6	0	–	1
6	–	1	6	0	–	–

Fig. 15-35 $S_{25,n41,31}$-(20).

1	2	3	4		
–	0	3	2	–	1
–	–	–	–	╳	–
–	11	–	0	–	0
–	1	0	–	–	9
10	–	6	0	–	2
13	1	6	0	–	–
10					

Fig. 15-36

1	–	0	3	2	–	1
2	–	–	–	–	X	–
3	–	11	–	0	–	0
4	–	1	0	–	–	9
5	0	–	6	0	–	2
6	3	1	6	0	–	–

Fig. 15-37 $S_{25,n41,n31}$-(29).

Fig. 15-38

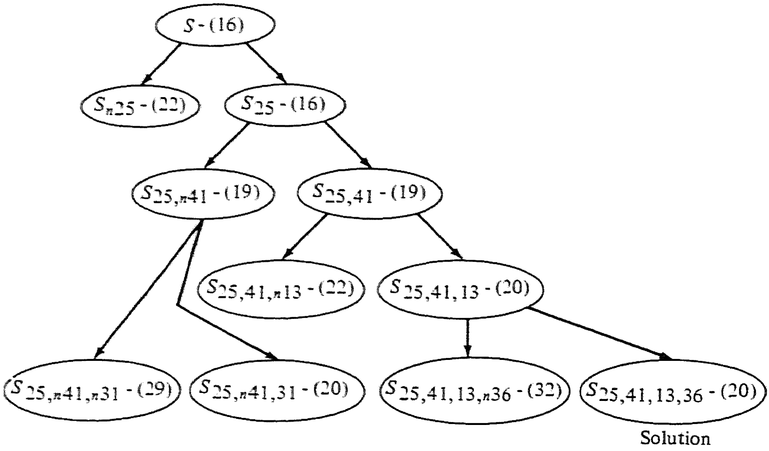


Fig. 15-39

solution shown in Fig. 15-38. Since all others have an equal or higher lower bound, we can conclude that this is the optimal solution.

Figure 15-39 summarizes the branching sequence which led to the solution. Although this procedure may have seemed laborious, it is much less so than complete enumeration. In this example complete enumeration would have required the evaluation of $n! = 6! = 720$ sequences.

n JOBS, $m = 2$ MACHINES, FLOW SHOP, STATIC ARRIVAL PATTERN, MINIMIZATION OF TOTAL ELAPSED TIME

Now we turn to a problem which can be categorized by n jobs, two machines, flow shop configuration, and a static arrival pattern. The criterion chosen will be the minimization of the total elapsed time for the entire set of jobs.

Johnson[4] showed that if the following set of rules was followed, t elapsed time would be minimized.

1. List the operation times on the two machines in two vertical columns.
2. Scan all the time periods for the shortest one.
3. If it is for the first machine, place the corresponding job first.
4. If it is for the second machine, place the corresponding job last.
5. Cross off both times for that job.
6. Repeat the steps on the reduced set of times. (The solution of the problem will therefore proceed from both ends toward the middle.)
7. In the case of a tie, order the job with the smallest subscript first. For example, if A_1 and B_4 are tied, select A_1 first. If the tie is between A_i, and B_i, order the job according to A.

Consider the example shown in Table 15-2. Here we have five jobs which must be processed through two machines. The jobs must pass through first A and then B. The problem is to find that sequence which minimizes total elapsed time. Following rule 1, the operation times for machines A and B are recorded in two columns. Then the columns are scanned for the shortest time. It is job 1 on machine B. Since it is associated with the second machine, the job is placed last in the sequence. The procedure is repeated for a reduced list of four jobs. The lowest value is 3 hours for job 2 on machine B. It is placed second to last. Next we find a tie between job 4 on machine A and job 5 on machine B. According to Rule 7, job 4 is sequenced first. We now have:

4			2	1

Continuing with the next selection, job 5 is placed third from the end and job 3 is entered in the only remaining slot left in the sequence. The optimal

---

[4] S. M. Johnson, "Optimal Two- and Three-stage Production Schedules with Setup Time Included," in John F. Muth and G. L. Thompson, *Industrial Scheduling*, Prentice-Hall, Inc., Englewood Cliffs, N.J., 1963, p. 13.

Table 15-2

Job i	Operation times: Machine A_i	Operation times: Machine B_i
1	~~3~~ hr	~~2~~ hr
2	6	3
3	9	8
4	~~4~~	~~6~~
5	7	4

sequence is

When the problem is extended beyond $M = 2$, it can no longer be solved by exact procedures. A heuristic presented by Campbell, Dudek, and Smith, however, finds nearly optimal solutions very efficiently. We now turn to their method.

n JOBS, m MACHINES, FLOW SHOP, STATIC ARRIVAL PATTERN, MINIMIZATION OF TOTAL ELAPSED TIME

The heuristic which we will employ for this n by m problem splits up the larger problem into a series of $n \times 2$ problems and then utilizes the rules generated in the last section to solve these $n \times 2$ problems. The best of them becomes the heuristic solution.

Given that there are m machines, we will generate $m - 1$, n-job, two-machine problems. They will be generated from the original problem in the following way. The first two-machine problem will contain the operation times for the first and last machines. For example, consider the five-job four-machine problem shown in Table 15-3. The first two-machine problem is recorded in Table 15-4. The second two-machine problem is derived by adding the opera-

Table 15-3

	Machine			
Job				
1	7	11	2	14
2	18	15	18	13
3	4	11	14	18
4	11	27	32	21
5	6	16	14	16

Table 15-4

	Machine	
Job	*A*	*B*
1	7	14
2	18	13
3	4	18
4	11	21
5	6	16

tion times on the first two and last two machines for each job. The result of this computation is shown in Table 15-5. The third two-machine problem is constructed by adding the operation times on the first three and the last three machines for each job. The results are shown in Table 15-6. Since there are $m = 4$ machines, and the number of two-machine problems required by this heuristic are $m - 1$ or $4 - 1 = 3$, we need not develop any more.

Using general notation, the operation times which we have just computed for the kth two-machine problem can be expressed as

$$P_{j1}^{k} = \sum_{i=1}^{k} t_{ji} \qquad j = 1, 2, 3, \ldots, n$$

$$P_{j2}^{k} = \sum_{i=m+1-k}^{m} t_{ji} \qquad j = 1, 2, 3, \ldots, n$$

where P_{j1}^{k} = operation time in kth problem for jth job on first machine
P_{j2}^{k} = operation time in kth problem for jth job on second machine
t_{ji} = operation time for jth job on ith machine

Now we apply the n-job, two-machine algorithm from the last section to the three auxiliary problems. The sequences which evolve from the application of this algorithm are shown in Table 15-7.

Table 15-5

	Machine	
Job	*A*	*B*
1	18	16
2	33	31
3	15	32
4	38	53
5	22	30

Table 15-6

	Machine	
Job	*A*	*B*
1	20	27
2	51	46
3	29	43
4	70	80
5	36	46

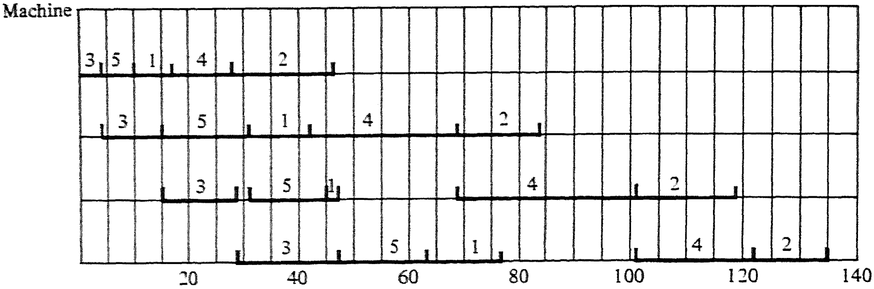


Fig. 15-40 Sequence 1: 3-5-1-4-2.

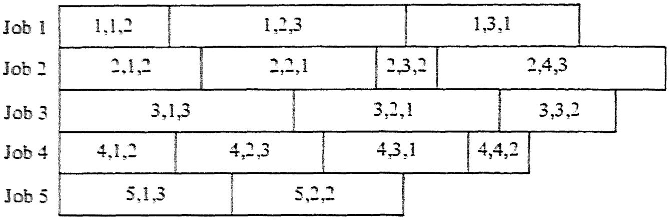


Fig. 15-41 Job scheduling data.

The final step is to compute the total elapsed time for each of the k sequences from the original $n \times m$ table of operation times. To accomplish this a Gantt chart must be drawn for each alternative sequence. A chart for the sequence 3-5-1-4-2 is shown in Fig. 15-40. It, in fact, turns out to be the best sequence. Elapsed time for the other sequences is shown in Table 15-7.

THE GENERAL n m JOB SHOP PROBLEM WITH A STATIC ARRIVAL PATTERN

We now turn to the general job shop problem in which we find n jobs to be processed on m machines in such a way that the routing for one job is not necessarily the same as the routing for the next. Consider the problem shown in Fig. 15-41, where the length of the rectangle represents the time necessary to complete the operation. The notation i, j, k represents the job number, the order in which that step must be performed, and the machine required for that

Table 15-7 Total elapsed time

kth auxiliary problem	*Sequence*	*Total elapsed time*
	3-5-1-4-2	135 hr
	3-5-4-2-1	138
	1-3-5-4-2	138

operation. The problem is to optimally rearrange these blocks into a schedule such that steps are performed in the proper order and that there is no overlap of jobs on a machine. One such feasible schedule is shown in Fig. 15-42.

Unfortunately there is no analytical solution to this general n/m job shop problem. The emphasis, moreover, has been on approximate solutions to the problem. One of these is merely to randomly generate a set of feasible schedules and choose from among them the one that best meets the objective. Other approaches include the use of heuristic rules to sequence jobs on machines. These include such rules as shortest processing time first, "first come, first served," and earliest due date first.

Before we take a closer look at these heuristics, one more assumption will be changed. In the preceding models we have assumed that all jobs arrived at one time. More common is the case where they arrive continuously. This is referred to as the *dynamic arrival* case.

THE GENERAL n/m JOB SHOP PROBLEM WITH DYNAMIC ARRIVAL PATTERN

When jobs arrive continuously in either a random or a continuous fashion, the problem of developing a schedule becomes even more complex. It is for this situation and the job shop with static arrival patterns that several priority rules have been tested. The rules are used to determine the next job that should be selected from a set of jobs awaiting processing on a particular machine. Some of these rules rank jobs according to:

1. Shortest operation time
2. First come, first served
3. Random order
4. Slack time remaining to due date (slack time = number of days to due date — number of processing days remaining)
5. Slack time per operation remaining
6. Length of next operation

The appropriateness of the rule, moreover, depends upon the criterion chosen. For example, a rule that minimizes the percentage of late orders may result in inefficient utilization of resources. Common criteria include the following:

1. Average flow time
2. Percent of late orders
3. Average waiting time of orders
4. In-process inventory costs
5. Labor utilization
6. Machine utilization

Making the choice of the priority rule even more difficult is the fact that the goals of the decision maker can very seldom be expressed in terms of a

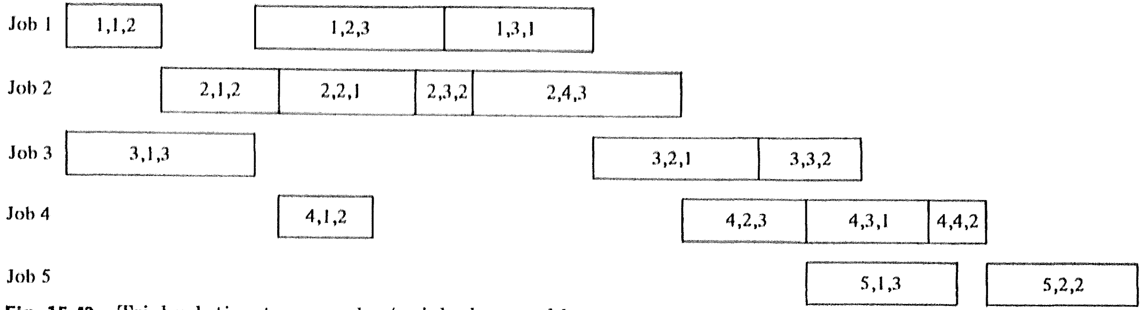


Fig. 15-42 Trial solution to general n/m job shop problem.

single criterion. Usually a combination of several factors must be considered in making a sequencing decision.

To test the effectiveness of these rules, large-scale simulations of job shop systems have been designed. Through these simulated shops, large numbers of jobs are processed according to the rule being tested. First, jobs are dispatched to the first machine in their specified order of operations; they then take their places in the queue based on the priority rule, wait for service, are processed by the machine, and finally are sent to the next machine. This pattern is continued until all the jobs have been processed or enough equivalent shop time has been simulated. During the simulation the important criteria by which the rule will be compared with others are recorded. Next, other rules are executed, and finally a comparison takes place between the performances of the various rules.

Several studies[5] have reached the conclusion that the "shortest operation time" rule has the lowest average flow time for all rules tested. That is, the average time spent processing the order will be the lowest if the jobs are scheduled on each machine with the highest priority going to the job with the shortest operation time.

Rather than focusing on the average processing time, it is more in line with industry practice to focus on the lateness of jobs. Slack time per operation, for example, has been reported to perform quite effectively in reducing the lateness of jobs. We now turn to a case study in which this priority rule was employed as a crucial part of a large-scale scheduling and control system.

A JOB SHOP SCHEDULING AND CONTROL SYSTEM

The scheduling and control of a job shop have characteristics which are common over a wide spectrum of applications. In general, the following sequence of events takes place. Upon finalization of an order, it is usually broken down into its basic parts or jobs. The fabrication plans for these jobs are sent to the appropriate locations in the shop, and the jobs are then scheduled at a work station or machine. If other jobs are already waiting, each job must take its place in the queue. The position which it takes within this queue depends upon the scheduling rule which has been chosen. It may be "first come, first served," shortest operation time, shortest setup time, and so on.

Once the order has been released to the shop, a control system usually monitors its progress. The primary purpose of this control system in many shops is to ensure that the due date will be met. When the order is not progressing according to plan, the job is expedited. In an unsophisticated system this might mean putting pressure on the foreman to keep the job moving, while in a more sophisticated system it might mean updating the job's relative priority and closely watching its progress through a computer-controlled system.

To give some substance to this descriptive framework, we turn to a scheduling and control system which was developed at Hughes Aircraft Company.[6] The job shop for which this system was designed fabricates a wide variety of products, including machined and sheet metal parts, machine assemblies, waveguides, and printed-circuit boards. At any one time there are between 2,000 and 3,000 orders being processed through the shop. The shop consists of 1,000 machines which are grouped into 120 functional work centers. Each work center is capable of performing one functional operation, such as milling, drilling, or grinding. The work centers are manned by 400 direct workers. Given this level of capacity, the average cycle time per completed job is three to four weeks.

---

[5] Elwood S. Buffa, *Production-Inventory Systems: Planning and Control,* Richard D. Irwin, Inc., Homewood, Ill., 1968, p. 304.

MAJOR FUNCTIONS OF THE CONTROL SYSTEM

The Hughes job shop's scheduling and control system performs four major functions. These include:

1 Preparation of a fabrication shop order and all necessary paperwork upon initiation of an order
2. Daily schedule for all parts in the shop
3. Daily order status through special reports
4. Weekly forecast of shop loads

Upon initiation of an order request, a *fabrication order* is prepared by data processing. Included in this packet are *master job cards,* which are eventually used together with plastic location cards to transmit status reports from the machine centers in the shop to a central information file.

The second function performed by the system is the determination of a daily schedule for each machine group. To sequence the jobs, a priority rule is used which minimizes job lateness. In general, whenever a job has fewer days until its promised delivery date, it takes on a higher priority than a job with more days. Jobs of higher priority are then scheduled first. Shortly we will take a closer look at the rule they use.

The scheduling problems would be somewhat simpler if it were satisfactory to schedule only those jobs which were to be found at the machine center each morning. But should we not take into consideration those jobs which are now at other machine centers and expected to arrive at the center being scheduled sometime during the day? In developing the daily schedule, neglecting these late arrivals would result in less than desirable schedules. A more comprehensive system will therefore schedule not only those jobs which are to be found at the machine center in the morning, but also those jobs which are expected to arrive at the center during the day. To accomplish this, an elaborate record-keeping system must be maintained to monitor the status of the job. In addition, a simulation must be performed whose output specifies the expected time of arrival and the updated priorities for these new arrivals. The completed schedule will therefore include jobs which are to be found in machine centers and jobs which are expected during the day. Data on the priorities for all jobs and the expected time of arrival for new jobs ensure a more effective sequence decision.

---

[6] Michael H. Bulkin, J. Colley, and H. Steinhoff, Jr., "Load Forecasting Priority Sequencing and Simulation in a Job Shop Control System," *Management Sciences,* vol. 13, no. 2, October 1966, p. B-29.

The third function of this system is to prepare daily status reports which can be used to control the jobs in process. Once the job is released to the shop, its progress is reported daily. A comparison can then be made between its planned and its actual progress, and appropriate expediting can be initiated if required.

The final function that this system performs is to forecast the shop loads for each machine group. The forecast is updated weekly and shows the load for the next 10 weeks. In this way, the shop utilization can be monitored and necessary short-term changes in capacity can be planned.

Next we turn to a closer look at the process of generating a daily schedule for each machine group.

PRIORITY RULE

For every job in the central information file a priority rating is computed. To accomplish this, two files must be set up and maintained. These include:

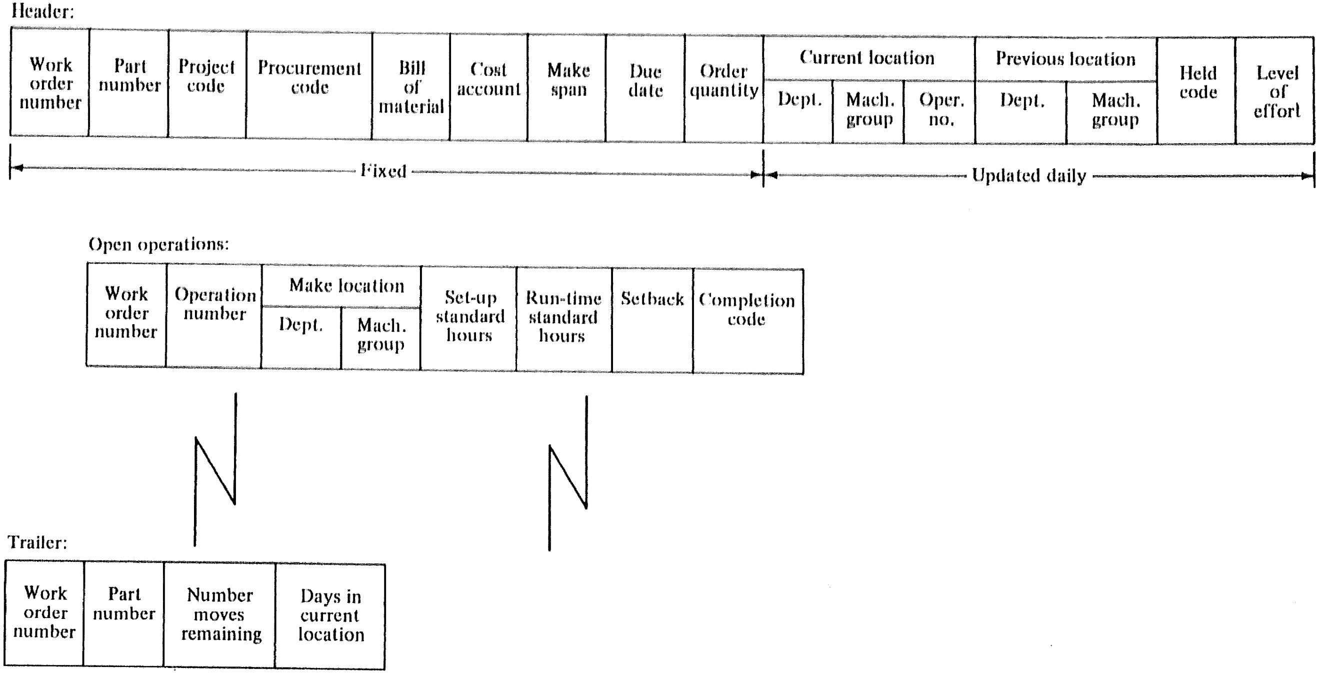
1. Planning and Standards Library
2. Fabrication Open-order Master File

The Planning and Standards Library includes setup and run time standards for parts built in the shop. Whenever changes occur, the file can be updated. The second file contains the information shown in Fig. 15-43. The second section of the Header record is updated daily by information received through remote terminals located in the shop. Here is where the status of the job is recorded. The next section of the record, Open Operations, specifies the sequence of steps required for job completion.

From this information, the priority of the job can be calculated. As we have seen, there are several rules which can be used to determine this priority. The particular rule chosen by this shop was the *minimum slack time per operation rule*. First, a slack time per operation is computed:

$$\text{Slack time per operation} = \frac{\text{RT} - \text{ROT}}{\text{NOR}}$$

where RT = remaining calendar time to due date
RRT = remaining operation time to order completion
NOR = number of operations remaining



Header:

Work order number	Part number	Project code	Procurement code	Bill of material	Cost account	Make span	Due date	Order quantity	Current location			Previous location		Held code	Level of effort
									Dept.	Mach. group	Oper. no.	Dept.	Mach. group		

Open operations:

Work order number	Operation number	Make location		Set-up standard hours	Run-time standard hours	Setback	Completion code
		Dept.	Mach. group				

Trailer:

Work order number	Part number	Number moves remaining	Days in current location

Fig. 15-43 Fabrication open-order master file. (*From Bulkin et al., "Load Forecasting Priority Sequencing and Simulation in a Job Shop Control System," Management Science, 13, 2 (October 1966), p. B-48.*)

Consider one particular job with 14 days left until its due date, 7 days of operation time, and three operations remaining. The priority can be computed in the following way:

$$\text{RT} = 14$$
$$\text{ROT} = 7$$
$$\text{NOR} = 3$$

$$\begin{array}{l}\text{Slack time}\\ \text{per operation}\end{array} = \frac{14-7}{3} = 3\tfrac{1}{3}$$

Since the remaining operation time does not include the time spent in a queue, the figure computed actually represents the clock time remaining per operation. Therefore, the *lower* this figure, the *higher* is the priority. In the event that the job is already beyond the scheduled due date, the slack time per operation would be negative. This would represent an even higher priority.

SIMULATION SCHEDULE

Before the daily shift begins, a priority is computed for each job in the Fabrication Open-order Master File. Then a shop simulation is run where these orders are assigned to machine centers and take places in the queue according to their priority ratings. The operation times are then retrieved from the file and each order is processed on the appropriate machine. After it is completed, the job is moved to its next location. Its position in this new queue depends upon its updated priority. The simulation for the entire shop continues for one shift.

The outcome of this simulation is summarized in the Order Schedule Report. An adaptation from it is shown in Table 15-8. It shows the orders in the machine center at the beginning of the shift and those which are expected to arrive during the shift. Also included in this report are the priorities for the jobs already at the station and the updated priorities for the incoming jobs.

The Order Schedule Report is the foreman's guide to scheduling his machine centers for the forthcoming shift. From it he knows what jobs he has and their priorities and what jobs are expected to arrive and their priorities and from where they are coming.

HOT ORDER VISIBILITY REPORT

A report that is quite useful for top priority orders is the Hot Order Visibility Report. This report shows where the hot orders are located at the beginning of the shift and their expected progress during the shift. Coordinators, whose responsibility it is to pay special attention to these orders, therefore know where each job is and when it is expected to move. Rather than pressuring the foreman

Table 15-8 Order schedule report†

Priority	*Part number*	*Work order*	*Previous location*	*Arrival time*	*Order quantity*	*Estimated hours of actual operation*	*Operations left*	*Estimated run time left*	*Date promised*
−64.1	17-3214	2241	12-07	—	170				
−10.2	14-3625	1731	12-07	—	210				
− .9	12-4168	3152	12-04	—	413				
.7	15-8764	1873	12-11	—	85				
1.3	10-1907	6249	13-01	—	170				
7.1	32-2165	8214	12-07	—	220				
14.2	18-4081	1988	15-02	—	100				
Incoming orders, Tuesday 6-14									
−13.2	16-3192	1753	12-09	10:20	160				
− 1.7	19-9162	2673	15-02	12:50	200				
3.1	81-1419	7175	13-01	1:30	350				
7.2	13-1217	6273	17-21	2:00	90				

Day	*Page*	*Machine group*
6-14	1	42-17

† Adapted from Bulkin, "Load Forecasting, Priority Sequencing, and Simulation in a Job Shop Control System," *Management Science*, vol. 13, no. 2, October 1966, p. B-35.

to process these jobs next, the coordinator is performing more of a control function.

RESULTS

Several significant improvements were reported as a result of this new system. The percentage of orders completed by their due date increased by 10 percent. The before-and-after profile of order completion is shown in Fig. 15-44. It is clear that the system has made an impact on delivery schedules.

This improvement in performance has led to the reduction of in-process inventory and a simultaneous increase in machine and manpower utilization.

Quite impressive is the reduction in expediting costs. These costs have been reduced by 60 percent, a savings large enough to pay for the computer costs of the simulation.

Case study: court scheduling[7] The scheduling of courtrooms is in drastic need of attention. Quite often cases are not heard on their scheduled dates because of a case overload. As a consequence of this, lawyers, policemen, and witnesses often waste an entire day in court. In fact, if the situation is repeated, it is likely that the witnesses will eventually refuse to try again.

Often the opposite occurs. A judge may close his courtroom early because of a lack of cases. Both of these problems can be attributed to a scheduling system which by present-day standards and capabilities can be considered antiquated.

Recently a study was undertaken for the New York City Criminal Court to see if a better scheduling system could be devised.

The consulting firm which designed the system identified several factors which guided their design effort. The first was that the probability of a judge running out of cases should be kept low. Second, the probability that a case would be adjourned because of a calendar overload should also be kept low. Third, a policeman's court appearances should be batched and scheduled for duty days. Fourth, high-priority cases should be scheduled as early as possible; and fifth, a maximum time limit should be enforced prior to which a case must be scheduled.

At the core of the scheduling system is a priority rule whose purpose it is to assign a priority rating to each case. Those cases with a high priority are slated for scheduling, while those of lower priority are returned to the unscheduled pool.

The priority assigned to each case depends upon several factors, including whether the defendant is in or out of jail, the seriousness of the charge, and the elapsed time since arraignment. Since priority is a function of time, the rating is continually being updated. When it reaches a threshold level, the case is scheduled.

Once a case is marked for scheduling, a multiple-regression routine

---

[7] This case is based upon the following paper: Samuel S. Shapiro, "An Automated Court Scheduling System," presented at the 12th American Meeting of the Institute of Management Sciences, Detroit, Mich., September 1971.

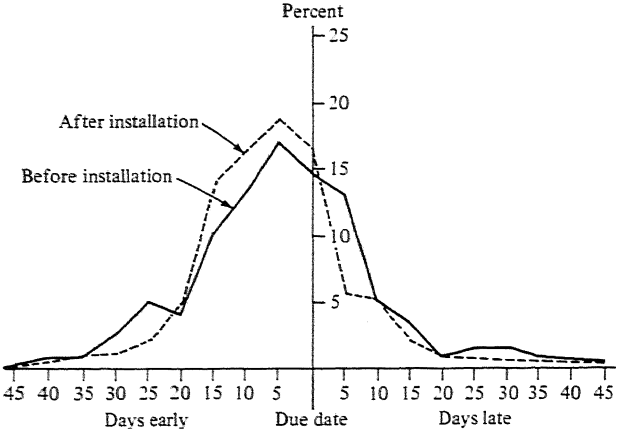


Fig. 15-44 Reduction in costs. (*From Bulkin et al., "Load Forecasting Priority Sequencing and Simulation in a Job Shop Control System," Management Science, 13, 2 (October 1966), p. B-47.*)

determines the time slot that would be necessary for the case to be heard. This model relates the expected court time to a set of causative factors such as seriousness of offense, plea of the defendant, number of witnesses, and presiding judge. Finally, the case is scheduled into an appropriate time slot.

The scheduling algorithm, however, does not schedule a full case load. It keeps time open for the insertion of high-priority cases. In fact, the further into the future the schedule extends, the more open time is reserved for priority cases.

Each day the schedule is revised. Cases which are new to the court and have received a high priority are added, as are the old cases whose updated priorities have exceeded the threshold.

To implement such a system requires a substantial investment in data collection, computer hardware, and additional system design effort. Furthermore it requires the approval and cooperation of lawyers, judges, and court administrators. At present the system is still considered to be in the design stage and none of these implementation problems have yet been confronted. The city is still open to design and implementation suggestions.

SUMMARY

It should now be quite apparent that the scheduling and control of an operational process can be a complex task. The appropriate methodology by which a solution to the problem is achieved should include not only a careful analysis of the operational process and its associated information flows, but also the design of efficient sequencing rules or algorithms.

For the very simple situations, such as n jobs through one machine, it becomes a relatively simple task to find the optimal sequence. However, as we progress to the general n-job-m-machine job shop problem, we have to be satisfied with solutions generated by heuristic rules.

For large-scale applications, a computer-based system is usually a must. Its primary functions would include the storage and collection of data, the generation of efficient schedules through the use of decision rules, and the provision of useful output for control purposes.

QUESTIONS

Q15-1. Why is a Gantt chart generally inadequate for machine-loading purposes?

Q15-2. Describe the sequencing problem.

Q15-3. If five people are waiting in a clinic for a series of tests to be made on them, how many different ways can they be sequenced through the process? Might it be advantageous to sequence them one way instead of another? Why?

Q15-4. Why would the type of facility (flow shop or job shop) affect the consequence of a sequencing strategy?

Q15-5. What makes the traveling salesman problem one of the most difficult in management science to solve?

Q15-6. In the job shop scheduling case, what decision rule was employed? Is it a heuristic or an optimization rule?

Q15-7. Given the complexity of the court scheduling problem, would you expect the computer to improve upon manual scheduling techniques now employed? What problems do you see in this system?

Q15-8. Is a computer a necessary component of a scheduling system? Explain.

Q15-9. Show that the area under the schedule graph in Fig. 15-3 represents the total processing time for that particular schedule.

Q15-10. Why is the technique which we have employed to solve the traveling salesman problem called a "branch-and-bound" algorithm?

Q15-11. What implementation problems would you foresee in the court scheduling case?

Q15-12. Why is the general n/m job shop problem so difficult to solve?

Q15-13. Suppose that you came up with a heuristic rule which you thought could perform quite well in sequencing jobs through a general n/m job shop facility. How would you test this rule and compare it with others?

PROBLEMS

P15-1. Solve the machine loading problem illustrated in Fig. 15-2.

P15-2. Eight jobs must be processed through a single machine. The operation time for each job is given below.

Job	*Operation time*
1	11
2	22
3	15
4	7
5	19
6	20
7	16
8	6

(*a*) What sequence will minimize the average processing time for the entire set of jobs?

(*b*) What sequence will minimize the average job lateness?

(*c*) What sequence will minimize the maximum lateness of any job in the set?

(*d*) What sequence will minimize the total processing time for all jobs?

P15-3. The length of time necessary to set up and run an ice cream manufacturing process depends upon the job previously run and the job about to be run. Given the following data, how should these four jobs be sequenced through the process?

To
job

From job				
1	–	5	9	2
2	7	–	5	11
3	4	8	–	6
4	3	7	12	–

P15-4. Solve the following traveling salesman problem.

To

From	1	2	3	4	5
1	–	3	11	9	12
2	7	–	3	6	4
3	14	5	–	8	13
4	6	9	4	–	8
5	8	5	1	10	–

P15-5. A salesman must make a tour of five cities. He must start and end in city 1. What route will minimize the distance traveled?

To

	1	2	3	4	5
	–	3	6	12	5
	4	–	5	5	4
	5	4	–	5	4
	9	5	6	–	9
	6	3	2	9	–

P15-6. Solve the following traveling salesman problem.

To

	1	2	3	4	5	6
1	–	7	10	11	6	7
	3	–	8	10	8	4
	4	5	–	9	3	11
	6	7	13	–	14	10
	8	4	5	2	–	7
	3	11	9	6	4	–

P15-7. Seven jobs must be processed through two machines. Each one must be processed by the first machine and next by the second. For the operation times which are given below, find a sequence such that the total elapsed time is minimized.

	Operation time	
Job	*Machine A*	*Machine B*
1	5	12
2	9	6
3	11	7
4	3	4
5	6	5
6	12	9
7	4	14

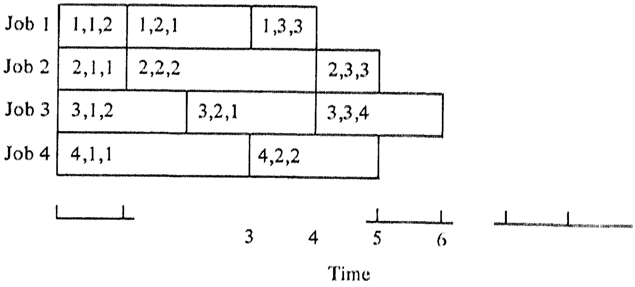
P15-8. Six jobs must be processed through three machines. Each job must first be processed on machine 1, then on machine 2, and finally on machine 3. For the operation times which are given below, find a sequence which achieves a low total elapsed time.

	Operation time		
Job	*Machine A*	*Machine B*	*Machine C*
1	12	6	9
2	7	9	13
3	8	11	5
4	4	6	11
5	5	13	6
6	10	5	8

P15-9. Three jobs must be processed on four machines. Each job must first be processed on machine 1, then on machine 2, next on machine 3, and finally on machine 4. Given the following operation times, find a schedule which comes closest to minimizing total elapsed time. Graph your results.

	Operation time			
Job	*Machine A*	*Machine B*	*Machine C*	*Machine D*
1	7	10	6	5
2	9	7	11	12
3	4	8	3	7

P15-10. Graphically illustrate a feasible solution to the following general n/m job shop problem.



BIBLIOGRAPHY

Buffa, Elwood S.: *Production-Inventory Systems: Planning and Control*, Richard D. Irwin, Inc., Homewood, Ill., 1968, chaps. 11 and 12.

Bulkin, Michael H., J. Colley, and H. Steinhoff, Jr.: "Load Forecasting, Priority Sequencing, and Simulation in a Job Shop Control System," *Management Science*, vol. 13, no. 2, October 1966, p. B-29.

Conway, Richard W., W. L. Maxwell, and L. W. Miller: *Theory of Scheduling*, Addison-Wesley Publishing Company, Inc., Reading, Mass., 1967.

Ferguson, Robert L., and Curtis H. Jones: "A Computer-Aided Decision System," *Management Science*, vol. 15, no. 10, June 1969, p. B-550.

Magee, John F., and David M. Boodman: *Production Planning and Inventory Control*, McGraw-Hill Book Company, New York, 1967.

Maxwell, William L.: "On Sequencing *n* Jobs on One Machine to Minimize the Number of Late Jobs," *Management Science*, vol. 16, no. 5, January 1970, p. 295.

Muth, John F., and Gerald L. Thompson: *Industrial Scheduling*, Prentice-Hall, Inc., Englewood Cliffs, N.J., 1963.

Niland, Powell: *Production Planning, Scheduling, and Inventory Control*, The Macmillan Company, New York, 1970, chap. 3.

Starr, Martin K.: *Systems Management of Operations*, Prentice-Hall, Inc., Englewood Cliffs, N.J., 1971, chaps. 10 and 11.

Wagner, Harvey M.: *Principles of Operations Research*, Prentice-Hall, Inc., Englewood Cliffs, N.J., 1969, chap. 13.

16

Quality Control

INTRODUCTION

It is early in the process design stage that conclusions on the attributes which must be incorporated in the product or service are finalized. These probably include decisions on size, weight, cost, maintainability, serviceability, performance, and appearance. Once determined, they serve as standards for system design and, later during the production cycle, as standards against which actual output must be measured. Good quality is therefore defined in relation to the standards set in the process design stage. If these standards are met, the quality is good; otherwise, it is not.

To ensure that these quality standards are met, control procedures are undertaken. Quality control can be exercised over the *product* or the *process*. When it focuses on the product, we call it *acceptance control.* Typically this involves making accept-or-reject decisions on a batch of parts *after* it has been through all or some of its operations. Usually this decision is based on a sample. Consider, for example, an order which has just arrived from a vendor. It contains 10,000 made-to-order insulators. The customer feels that a decision on whether to accept or reject this lot can be based on a sample. If the results from the sample are satisfactory, the lot will be accepted; otherwise, the lot will be sent back to the vendor.

Quality control is also used to ensure that processes are operating under control. Consider a lathe which has been set to machine a piece of aluminum to $1\frac{1}{2}$ inches in diameter. To verify this setting, a random selection of pieces could be periodically selected and measured. In this way, whenever the machine goes out of control, remedial action could be taken *before* too many defective items are manufactured. This aspect of quality control is called *process control.*

The approach which will be taken in this chapter is to first consider the appropriate places in the transformation process where quality control measures should be undertaken. Then we will turn our attention to the concepts of sampling risk and the economics of sampling. Following this the design of acceptance control and then process control systems is considered. A case is then presented which explores the application of these concepts to bank auditing. In the last, but optional, section bayesian sampling techniques are illustrated. These techniques allow the decision maker to explicitly incorporate his prior beliefs on process quality into the quality control decision process.

WHERE TO INSPECT

An important, but usually neglected, aspect of the quality control problem is the decision where to inspect. In Fig. 16-1, a three-stage production sequence is shown. Inspection can take place after production has been completed or any place between production stages. If inspection was costless, we might inspect at all of these potential locations. Another reason we might inspect at every point would be a very high cost associated with passing defectives through the process.

Consider the manufacture of a precision piece of navigational equipment to be used on the Lunar Excursion Module. At each successive production stage, a costly set of operations is performed on the unit. If at any stage the unit becomes defective, any further processing will just add costly steps to a now worthless unit. It is, therefore, reasonable to accept the moderate cost of thorough inspection after every stage rather than the high expected cost of passing defectives through the system. For most products, however, one would not expect inspection to take place after every production stage.

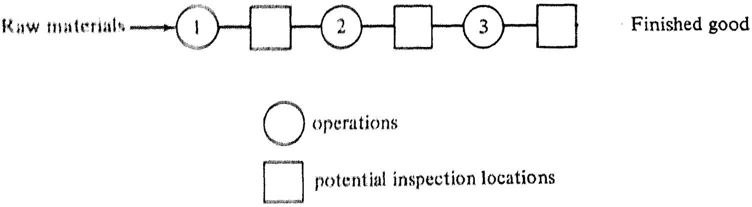


Fig. 16-1 Location of inspection stations.

In general, the number and location of inspection stations will depend upon two costs. The first is the inspection cost; this will *increase* as the number of inspection locations increase. The second is the cost of passing on defectives through the system. As the number of inspection locations increases, this cost will *decrease*. The optimum number and location for these inspection stations are achieved when the sum of these two costs is a minimum.

There is no simple procedure for solving this problem. Heuristics have been used; however, a digression into these methods is beyond the scope of this chapter. Suffice it to say that the number and location of inspection stations are an important consideration in the design of the overall quality control system.

SAMPLING RISK AND THE ECONOMICS OF SAMPLING

SAMPLING RISK

Once the decision is made to inspect at a particular stage, the next step is to develop a sampling plan. Consider a lot of machined parts which has just been received from the vendor. Before they can proceed through the manufacturing process, it must be determined whether or not this lot meets the desired standards. An acceptable lot is defined as one with 5 percent or less defectives. One way to establish its quality is to examine the lot completely—that is, to use 100 percent inspection. This strategy will ensure that no lots in excess of 5 percent defectives will be passed on through the manufacturing process.

Another strategy would be to sample a small portion of the lot and accept or reject it on the basis of the sample results. Consider that the lot of machined parts contains 10,000 pieces and that a sample of 100 is taken. Assume that the decision rule is to accept the lot if the number of defectives in the sample is less than or equal to five. Suppose that, unknown to us, the lot actually has 6 percent defectives, or 600 defective pieces. It is quite possible that for any sample of 100 pieces, we could be "unlucky" enough to draw four defective ones. Our decision rule, however, would tell us to accept the lot. What we have, in fact, done is to accept a bad lot. This is the first type of risk that you take when sampling; it is called *the consumer's risk*. How can one reduce this risk? Take a larger sample. The limit, of course, would be a 100 percent sample, for which no risk is taken.

There is still another kind of risk that is incurred. Suppose that our next lot has, unknown to us, an actual percent defective of 3 percent. It is quite possible for us to take a sample of 100 pieces in which six defective items are found. Following our decision rule we would reject this lot. What we have done this time is to reject a good lot. The likelihood of this happening is called *the producer's risk*. Again the larger the sample size, the smaller is this risk.

THE ECONOMICS OF SAMPLING

We have seen that whenever a sample is taken, risks are incurred. First, bad lots can be sent on through the manufacturing process. The consequence of this is the cost of passing these defectives through the process. Bad lots which are accepted by the inspection process and sent to further manufacturing stages or on to the customer before being discovered incur costs which could have been avoided by 100 percent inspection. Secondly, perfectly good lots can be rejected. The consequence of this is the unnecessary rework and administrative costs that are incurred.

Two kinds of costs are therefore born at an inspection location. The first is the cost of sampling and the second is the expected cost associated with sampling risk. The behavior of these costs is such that as the sample size increases, inspection costs rise but the costs associated with sampling risk fall. The optimal strategy is that level of inspection for which the sum of these two costs is a minimum.

Theoretically, then, a sampling plan should be based on estimates of these two costs. Practically, however, it is seldom done. The reason is that they are very difficult costs to assess, especially the cost of passing on defectives. For example, a judgment of the cost of passing defectives at the final stage of inspection must consider those direct costs borne by the firm when the customer receives a defective unit as well as those indirect costs of the loss of customer goodwill.

When we turn to the development of a sampling plan shortly, you will see that consideration of these two costs will not be an explicit part of the analysis. It will be shown, however, that the level of these costs is probably implied and that their omission in the first place is in the interest of practicability. We now turn to a more formal way of measuring sampling risk.

ACCEPTANCE SAMPLING BY ATTRIBUTES

First we will develop a statistical sampling plan where the results of each item inspected can be classified as either acceptable or defective. That is, either the part is good or it is bad. This is called *acceptance sampling by attributes*. An example of this is the inspection of a machined shaft where a set of go and no-go gauges might be used to accept or reject the part. The first gauge is used to test for the minimum shaft diameter. If the shaft fits through the test hole, this means that the shaft is undersized, and it is recorded as a defect. The second gauge is used to test for oversized shafts. If the shaft fails to fit through the gauge, it is rejected as being oversized. In this way the process is continued until all of the items in the sample are inspected and categorized as either good or bad. Then it is a simple matter to compute the sample lot's percent defective.

The first step in designing such a sampling plan is to determine what constitutes acceptable quality for the lot. Is it 5 percent defectives or less, 2 percent defectives or less, or 10 percent defectives or less? The answer depends, of course, upon the product.

Whatever this level, we would like to accept all lots whose actual percent defective is less than or equal to our acceptable quality level (AQL) and reject all others. The only plan which can do this perfectly, however, is a 100 percent sampling plan.

OPERATING CHARACTERISTIC CURVE

On graphs like that shown in Fig. 16-2, the discriminating power of any sampling plan can be illustrated. Along the abscissa is measured the actual percent defective in the lot, and along the ordinate is measured the probability that a lot with a certain actual percent defective would be accepted.

A curve that shows the relationship between the probability of acceptance and the actual percent defective is called an *operating characteristic* (OC) *curve*. The OC curve for a 100 percent sampling plan is shown in Fig. 16-2. You can see that this plan perfectly discriminates between good and bad lots. For example, if the actual percent defective in the lot were less than 5 percent, the lot would be accepted ($P_a = 1.0$). If, on the other hand, the actual percent defective exceeded 5 percent, the lot would be rejected with certainty ($P_a = 0.0$). Clearly this plan is riskless. Now let us examine the discriminating power of a smaller-size sample.

A sampling plan must include two pieces of information. The first is the size of the sample n that must be drawn from the lot. The second is the maximum number of defectives c that can be uncovered in the sample before the lot is rejected. Consider a sampling plan with $n = 100$ and $c = 5$. From the lot of 10,000 pieces we will take a random sample of 100 pieces and test them to determine if they are good or bad. If we find *more* than five defective pieces, the lot will be rejected. The discriminatory power of this plan can be seen in Fig. 16-3. For example, if the actual percent defective of the lot were 3 percent, this sampling plan would accept the lot 92 percent of the time. Therefore, 8 percent of the time lots of this high quality would be rejected.

Consider a lot whose actual percent defective is in fact at the acceptable quality level, 5 percent. This plan would accept lots of this quality only 62 percent of the time. Consequently 38 percent of the time lots of this quality would be rejected. This represents the producer's risk. If the lot were as bad as 10 percent actual defective, this plan would accept the lot 7 percent of the time.

As you can now see, there is not perfect discrimination around the AQL level. It can be improved, however, by increasing the sample size. In Fig. 16-4, for example, another plan with $n = 200$ and $c = 10$ is presented, together

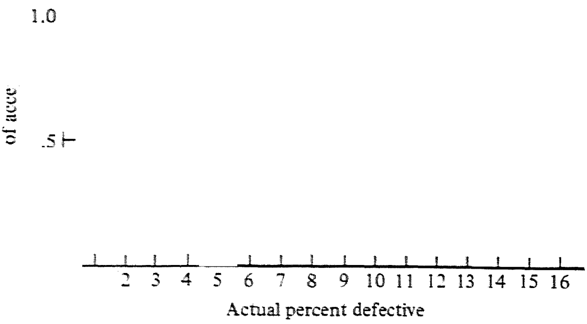


Fig. 16-2 100 percent sampling plan.

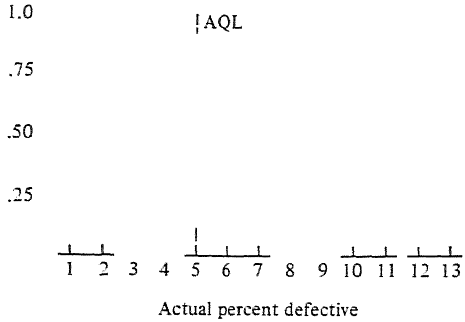


Fig. 16-3 Sampling plan: $n = 100; c = 5$.

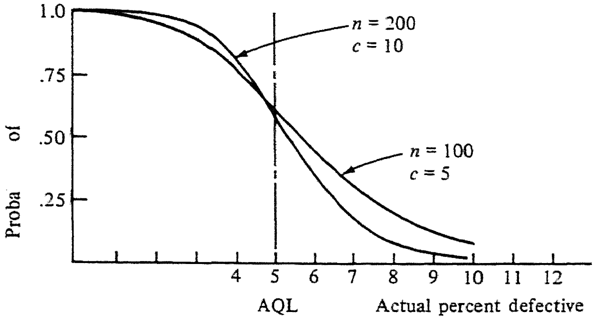


Fig. 16-4 Comparison of sampling plans.

with the old one. The discriminatory power of the new plan is clearly better than that of the first. The offsetting factor, however, is the higher cost of the larger sample.

Drawing the OC curve Whenever the lot size is large and the ratio of the sample size to lot size is less than .20, the cumulative Poisson distribution can be used to approximate the probability of acceptance for any level of percent defectives. A cumulative Poisson chart, for this purpose, can be found in Appendix B.

Suppose that we are interested in constructing the OC curve for the plan specified in Fig. 16-3. We will proceed by computing the probability of acceptance for several levels of actual percent defective. Then an OC curve will be drawn through these points. For example, the probability of acceptance for 1 percent actual defectives can be found in the following way. First the expected number of defectives is computed:

$$u = \frac{n(\text{PD})}{100}$$

where u = expected number of defectives
n = sample size
PD = actual percent defective

In this case

$$u = 100 \; {}_{100} = 1$$

We then enter the cumulative Poisson table down the first column until this value of u is found. Now we read across to the column which is headed with $c = 5$. The figure that we find, .999, is the probability of *c or less* defectives in a sample which has an average number of defectives equal to u. Since whenever c or fewer defects are found, the lot will be accepted, we can label this the probability of accepting the lot.

Next we proceed to find the probability of acceptance when the percent defective is 2 percent. We have

$$u = 100 \frac{2}{100} = 2$$

and reading across to $c = 5$, we find the probability of acceptance to be .983. We continue in a similar fashion, find the remaining points, and then plot them on a graph. The results are shown in Table 16-1 and Fig. 16-3.

Now that we have learned how to construct an OC curve *given* a certain sampling plan, we turn to the problem of *finding* a sampling plan that will achieve a specified degree of discrimination.

THE SAMPLING PLAN

Discrimination specified by α, β, AQL, and LTPD The purpose in the design of a sampling plan is to choose a combination of n and c such that lots of high quality will be accepted with a high probability. We have specified this high level of quality as AQL, and to be more explicit we will also identify the high probability as α. For example, we might wish to find a sampling plan which has an $\alpha = 95$ percent chance of accepting a lot of AQL = 3 percent defectives.

Table 16-1 Determining the probability of acceptance

Actual percent defectives	$u = \dfrac{n(PD)}{100}$	c	*Probability of acceptance*
1	1	5	.999
2	2	5	.983
3	3	5	.916
4	4	5	.785
5	5	5	.616
6	6	5	.446
7	7	5	.301
8	8	5	.191
9	9	5	.116
10	10	5	.067

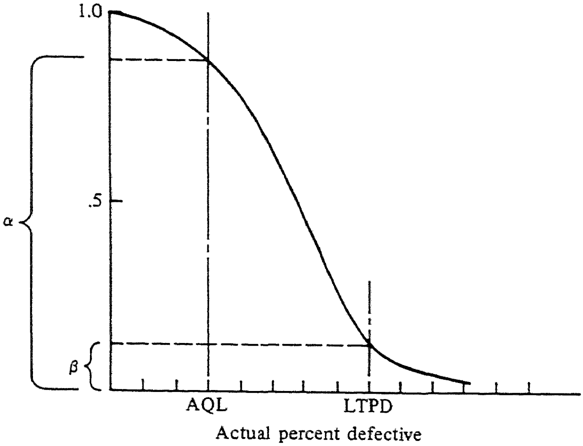


Fig. 16-5 Complete specification of OC curve.

The values of AQL and α do not, however, completely express the degree of discrimination required. The OC curves of Figs. 16-3 and 16-4 also illustrate that there is a chance of accepting lots of poor quality (consumer's risk). We can define the level of poor quality as LTPD (lot tolerance percent defective) and the probability that lots of this quality will be accepted as β percent. Of course we would like to have a low probability of accepting this poor quality. Now we can specify two points which summarize the degree of discrimination desired. In Fig. 16-5 the way in which these four values, including AQL and α, together with LTPD and β, completely specify an OC curve is shown. Once these values are decided upon, the next step is to find that sampling plan, n and c, whose OC curve passes through these points.

The first step, then, is to specify the values for AQL, α, LTPD, and β. It is quite common for α to be set equal to .95 and β equal to 10 percent. Suppose the AQL for a particular part is set at 3 percent, α at .95, β at .10, and LTPD at 11 percent. Now the problem is to find a sampling plan which achieves this degree of discrimination.

A computational procedure for determining a sampling plan In Table 16-2, a procedure for finding the values of n and c is illustrated for this problem. In the first column is found a sequence of possible values of c. The second column represents the values of $n(\text{PD}_{\text{AQL}})/100$ which can be read from the cumulative Poisson chart. For example, consider the first value for $c = 0$. The cumulative Poisson chart is entered in the $c = 0$ column, and we read down until we arrive at probability of acceptance P_a associated with AQL; this is .95. Reading across, we find $n(\text{PD}_{\text{AQL}})/100$ is approximately .05. Next we enter the $c = 1$ column and read down to .95. We find $n(\text{PD}_{\text{AQL}})/100 = .35$. The process continues in a similar fashion.

Column 3 represents the values of $n(\text{PD}_{\text{LTPD}})/100$. This time, when read-

Table 16-2 Determination of sampling plan

(1) c	(2) $\alpha = .95$ $\frac{n(PD_{AQL})}{100}$	(3) $\beta = .10$ $\frac{n(PD_{LPTD})}{100}$	(4) $\frac{(3)}{(2)}$
0	.05	2.30	46
1	.35	3.90	11
2	.85	5.30	6.2
3	1.40	6.70	4.8
4	2.00	8.00	4.0
5	2.60	9.40	3.6

ing down the columns of the cumulative Poisson table, we search for $P_a = .10$. In the first case, for $c = 0$, we find that $n(\text{PD}_{\text{LTPD}})/100 = 2.30$. Again the process continues for the other values of c.

Let us digress for a minute so that the theory behind column 4 will be understood. The ratio of column 3 to column 2 can be simplified in the following way:

$$\frac{n(\text{PD}_{\text{LTPD}})/100}{n(\text{PD}_{\text{AQL}})/100} = \frac{\text{PD}_{\text{LTPD}}}{\text{PD}_{\text{AQL}}}$$

We can compute this ratio from the data which were initially given:

$$\frac{\text{PD}_{\text{LTPD}}}{\text{PD}_{\text{AQL}}} = \frac{11}{3} = 3\tfrac{2}{3}$$

Therefore the appropriate value of c is the one whose corresponding ratio of column 3 to column 2 is closest to $3\frac{2}{3}$. In column 4 this ratio is computed, and it can be seen that at $c = 5$ the ratio is closest. Note that the ratio of column 3 to column 2 is continually decreasing. It is, therefore, unnecessary to look further than $c = 5$ in the hopes of finding a closer approximation.

Now that c has been determined, we turn to n. Solving for n can be accomplished by using either column 2 or column 3. The answer will not be the same for each one, but it will be close. We will do it both ways, and the reason behind these differences will be pointed out.

Using the second column, we have

$$\frac{n(\text{PD}_{\text{AQL}})}{100} = 2.60$$

Rearranging terms, we have

$$n = \frac{2.60}{\text{PD}_{\text{AQL}}/100}$$

Since $PD_{AQL} = 3$, $n = 2.60/.03 = 87$. Therefore the sampling plan which we have developed is

$$n = 87$$
$$c = 5$$

Let us check the α and β probabilities at AQL and LTPD. At AQL we have $n(PD_{AQL})/100 = 87(.03) = 2.6$. From the cumulative Poisson table we find that $P_a = .95$. We can conclude that our OC curve goes right through AQL = 3 percent and $\alpha = .95$. But this is just as you would expect, since we have verified the P_a at AQL by working backward through our solution for n and c.

Checking on the P_a at LTPD, we find a different story. First we compute $n(PD_{LTPD})/100 = 87(.11) = 9.57$. From the cumulative Poisson table we can read that for a value of $c = 5$ and $n(PD_{LTPD})/100 = 9.57$ the $P_a = .085$. That is, *by using column 2* to determine n, the α error is exact but the β error is not; the OC has *almost* gone through our two points but misses at LTPD and β by just a little. We therefore say that α is fixed but β floats if we solve for n using the second column.

What would have happened if the third column were used to solve for n? The answer is that β would be fixed and α would float. To illustrate this, our problem is solved using column 3. We find the following:

$$n = \frac{9.40}{.11} = 85$$

with a sampling plan of

$$n = 85$$
$$c = 5$$

Checking on the value of α, we find

$$\frac{n(PD_{AQL})}{100} = 85(.03) = 2.55$$

and from the cumulative Poisson chart we find that P_a at AQL is .958.

We can therefore conclude that the combination of n and c will depend upon whether α is fixed and β floats or α floats and β is fixed. In other words, it is usually not possible to find a plan whose OC curve goes through both points, thereby achieving the precise degree of discrimination desired. Which of these two plans to finally select depends upon which one comes closest to meeting the designer's needs. For example, it might be more important that α be fixed and β float than the other way around. Usually, however, the difference between these two plans will be negligible.

The appropriate degree of discrimination Now that the process of designing a sampling plan is familiar, we return to the problem of establishing the appropriate degree of discrimination. Suppose it was decided that the plan just developed did not provide enough discrimination. A new plan could be designed based on the following criteria.

$$\text{AQL} = 3\% \qquad \text{LTPD} = 7\%$$
$$\alpha = .95 \qquad \beta = .10$$

Since LTPD has been moved much closer to AQL, it is clear that the new design criteria are much more discriminatory. The consequence of this new plan will be a much higher sample size n. Since larger sample sizes are more costly to undertake, the cost of this new plan would be higher than that of the original one. The benefit, however, is that the likelihood of passing defective lots is reduced. Therefore, we have paid a higher cost to reduce the risk.

It is this trade-off between reduction in risk and increase in sampling costs which must be carefully considered when the initial design criteria are specified. Since the economic consequence of the reduction in risk is difficult to assess, frequently this aspect of the design problem is not given the appropriate attention.

Average outgoing quality As a consequence of processing a lot through a sampling stage, the outgoing quality will be as good as or better than the incoming quality. Lots which have a small number of defective items will be passed with a high probability, and the outgoing lot will therefore be of high quality. Lots which have a large number of defects will be rejected with a high probability, will usually then be subjected to 100 percent inspection, only good parts will be passed on, and as a result the average outgoing quality will again be high.

This phenomenon, for the plan $c = 5$ and $n = 87$, can be seen in Fig. 16-6. Along the abscissa is the incoming quality and on the ordinate is the average outgoing quality (AOQ). It can be seen that the worst that the average outgoing quality can be for this plan occurs when the incoming quality is about 5 percent. At that point the average outgoing quality is 3.6 percent; this is called the *average outgoing quality limit* (AOQL).

DOUBLE AND SEQUENTIAL SAMPLING

Double sampling There are economic advantages to be gained when sampling strategies more sophisticated than those which we have been using are employed. One such strategy is double sampling. In a double sampling plan, first a sample, of size n_1, is randomly selected from the lot and inspected. If it contains less than c_1 defectives, it is accepted; and if it contains more than c_2 defectives, it is rejected. If, however, it contains between c_1 and c_2 defectives, another sample, of size n_2, is taken. If the combined sample of $n_1 + n_2$ contains less than c_2 defectives, the lot is accepted; otherwise it is rejected.

The value in double sampling is that the first sample size n is much smaller than what would have been taken in a conventional single sampling

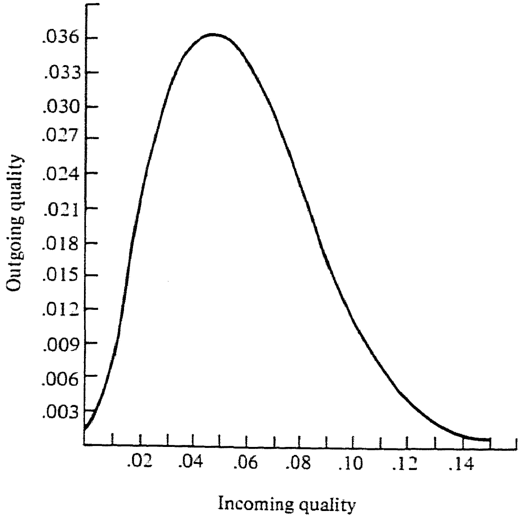


Fig. 16-6 Average outgoing quality.

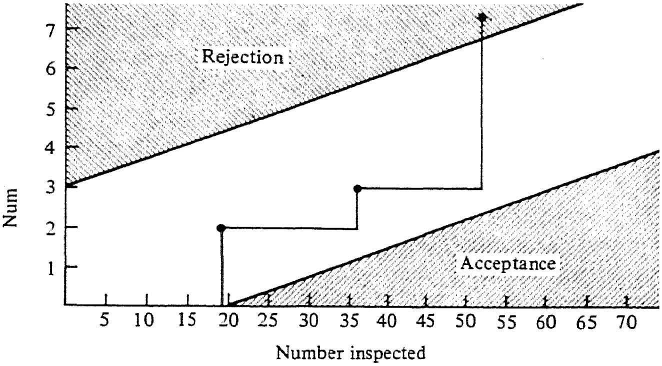


Fig. 16-7 Sequential sampling.

plan. Therefore, if the lot is good enough to be accepted or bad enough to be rejected on the smaller sample size, lower sampling costs are incurred.

Sequential sampling Sequential sampling is but a logical extension of double sampling. The difference is that decisions are made after the inspection of a *sequence* of small samples. In fact, the sample size may be as small as 1. After each sample is drawn, the items are inspected. The cumulative result of this and the previous samples is then used as the basis to either accept the lot, reject the lot, or take another sample.

An illustration of how the sequential sampling plan works is shown in Fig. 16-7. If the cumulative number of rejects remains outside of the shaded regions, sampling must be continued. If, however, the results fall in the shaded regions, the lot is accepted or rejected.

ACCEPTANCE SAMPLING BY VARIABLES

The sampling plans that have been studied so far were based on sampling techniques which determined whether a part was either good or bad; no actual measurements were recorded. There is, however, another way to inspect. Rather than just making go or no-go decisions, actual measurements can be taken. Such a plan is called a *variable sampling plan*. Consider the following variable sampling plan for ensuring that the shaft diameters of a particular lot are not too large. Let $n = 15$ and $c = 1.75$. First, 15 shafts are selected at random from the lot. Next the diameters are measured, and then the average diameter is computed. The result of one such sample of 15 is a sample mean of 1.33 inches. Since this is less than $c = 1.57$, the lot is accepted. Had it been greater than 1.57, the lot would have been rejected.

The criteria for the design of this sampling plan are exactly the same as in attribute sampling: AQL, α, LTPD, and β. The process of going from there to the determination of n and c, however, is quite different. Rather than using distributions which focus on proportions, we use the normal distribution. The detail of generating such a plan, however, is beyond the scope of this chapter.

It is usually the case that to take actual measurements is much more time-consuming than to make a go or no-go decision. Therefore, we can say that the inspection process for variable sampling plans is more costly than it is for attribute plans. The benefit, however, is that variable plans require smaller samples to achieve the same level of protection as attribute plans. For example, the same degree of protection can be achieved by an attribute sampling plan in which $n = 300$ and by a variable sampling plan where $n = 85$.

In some cases where inspection destroys the part, variable sampling is certainly superior. For example, in testing light bulbs, a sample set of bulbs is illuminated until each one fails. The average life is computed and compared with the critical value c. If the actual average is greater than the critical value, the lot is accepted; otherwise it is rejected.

In other cases where it might not be clear whether the sampling plan should be by variables or by attributes, the total cost of sampling and measuring must be evaluated for each alternative. The approach with the lowest total cost is then chosen.

PROCESS CONTROL

Now we turn to the problem of monitoring a process in order to determine whether or not it is operating in control. As long as the output is satisfactory, we can conclude that the process is "in-control." If, however, the output is not satisfactory, the process is "out-of-control," and appropriate corrective action should be taken. In this way the manufacture of lots with a high percentage of rejects can be prevented.

The output from any process is subject to two kinds of variability. These include variability which can be attributed to both chance and assignable causes. Chance variations are caused by many factors, none of which accounts for much variation by itself. Some of these sources for a machine might include changes in electrical voltage, wear of gears, changes in temperature, uneven tool wear, and so on. The consequence of several sources of error is that they occur in random fashion, and there is little that can be done to eliminate them.

There frequently occur, however, variations which can be assigned to particular causes. These include differences between workers, differences between machines, differences between materials, and differences due to a combination of these factors.

When chance variations are the only source of variation, the process is said to be in-control. However, when variation is due to one of these assignable causes, the process is out-of-control. The problem is to determine when the process is out-of-control, so that remedial action can be taken. The reason why this is a problem is that both chance and assignable variations occur simultaneously. The solution is to set upper and lower control limits. Variations within these limits will be assumed to be from chance, and no remedial action will be taken. Variations outside of these limits will be attributed to assignable causes and remedial action will be initiated.

CONTROL CHARTS FOR VARIABLES

Monitoring process mean Consider a machine which has been adjusted to produce shafts 1.50 inches in diameter. Experience has shown that the standard deviation of the process, when it is in-control, is .21 inch. The problem is to design a control system for this process.

Usually control systems are designed for a sample of several items from the process rather than for individual measurements. This is primarily intended to reduce sampling risk but also makes designing the control system much easier, since the distribution of the means of samples is approximately normal for large samples, whereas the distribution of individual measurements may not be.

Suppose that we elect to take a sample of $n = 10$ shafts at random from the machine. The distribution of sample means for all possible samples of size $n = 10$ is shown in Fig. 16-8. This is called the *sampling distribution of the mean,* and it is quite different from the distribution of the individual values

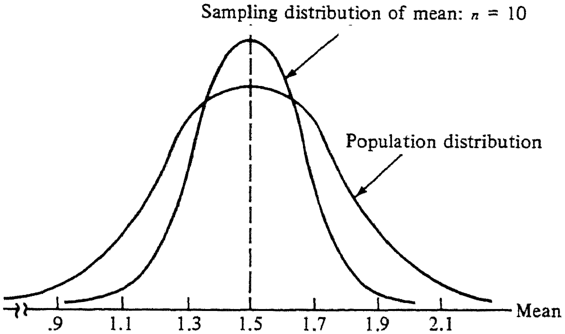


Fig. 16-8 Population and sampling distribution of mean.

in the population also shown in Fig. 16-8. To make this distinction clearer, consider the following example. From our knowledge of the standard deviation and the assumption that the output from the process is normally distributed about its mean $\mu = 1.50$, we could determine the likelihood that any one shaft would be greater than, say, 1.92 inches. This requires that we find the area to the right of 1.92 inches. To enter the normal table we need to express the value of 1.92 in terms of standard deviations from its mean; it is $(1.92 - 1.50)/.21 = 2$ standard deviations. Entering the normal table at 2, we find that the area above this value is .0228. That is, the likelihood that any *one* shaft which has been taken from this process will exceed 1.92 inches is 2.28 percent.

Suppose that instead of taking one item from the population, a sample of 10 items was taken. The likelihood that the mean from these 10 items would exceed 1.92 is much smaller than 2.28 percent. Why? Because now the relevant variability of outcomes is not the variability of the population, but the variability of sample means of size 10 as summarized by the sampling distribution of these means. The variability of the sample means is much smaller than that of the population's because in a sample of 10, mean values at either extreme are especially rare. This can be attributed to the fact that large values are likely to be offset by small values. The result is that the sample means will cluster quite closely around the mean of the population.

The variability of a sampling distribution of means, also called the *standard error of the mean*, for a sample of size n can be approximated by

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$

where σ = population standard deviation
n = sample size

Since σ is seldom known, an estimate utilizing the standard deviation computed from a sample is frequently used.

Returning to our example, we have the following:

$$\mu = 1.50$$

$$\sigma = .21$$

$$\frac{.21}{\sqrt{10}} = .07$$

Using the normal table found in Appendix B, we can make the following statements:

1. Ninety-nine percent of the time, chance variations of the mean will be contained within 1.50 ± 2.58(.07), or from 1.32 to 1.68.
2. Ninety-five percent of the time, chance variations of the mean will be contained within 1.50 ± 1.96(.07), or from 1.36 to 1.64.

3. Sixty-eight percent of the time, chance variations of the mean will be contained within $1.50 \pm 1(.07)$, or from 1.43 to 1.57.

Suppose that a control system was designed such that the upper control limit (UCL) was set at $1.50 + 2.58(.07) = 1.68$ and the lower control limit (LCL) at $1.50 - 2.58(.07) = 1.32$. These upper and lower control limits are used in the following way. Whenever the outcome of a sample of size 10 falls outside these limits, the machine is considered out-of-control. If, however, the result is within these limits, the process is assumed to be in-control.

As is true in any sampling process, risks are incurred, and this time is no exception. If the upper control limit is set at 1.68 and the lower at 1.32, 1 percent of the time chance variations from the process will exceed these limits (producer's risk). When this happens, remedial action will be inappropriately taken. This is called a Type I error, one in which action is taken when none is actually required. The second kind of error occurs when a variation could be traced to an assignable cause but the variation is not large enough to exceed the control limits. In this case, no action is taken when some, in fact, is required; this is referred to as a Type II error.

The nature of the risk taken can be altered by changing the UCL and LCL levels. For example, if these levels are narrowed to 1.36 and 1.64, the chance of a Type I error increases and the chance of a Type II error decreases. The best plan is the one which minimizes the cost of these two errors.

Example The usual procedure followed in a statistical control program is to plot the sequence of samples, as they are taken, on a control chart. A control chart for 25 consecutive samples is shown in Fig. 16-9. It appears that the process is in-control.

Monitoring process variability The control charts which we have discussed so far have monitored the mean value of the process. When it is suspected that

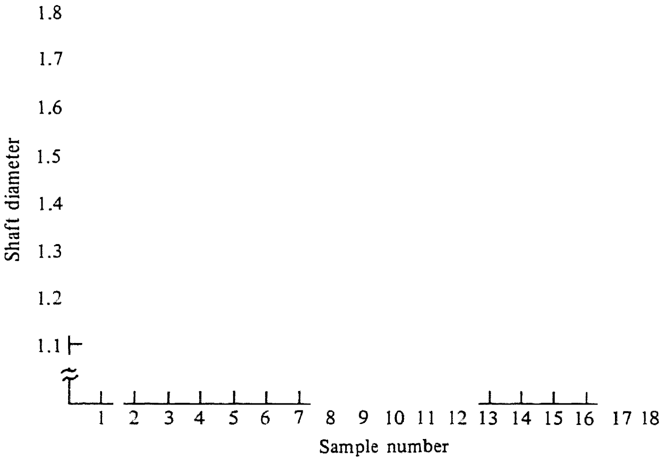


Fig. 16-9 Control chart for process mean.

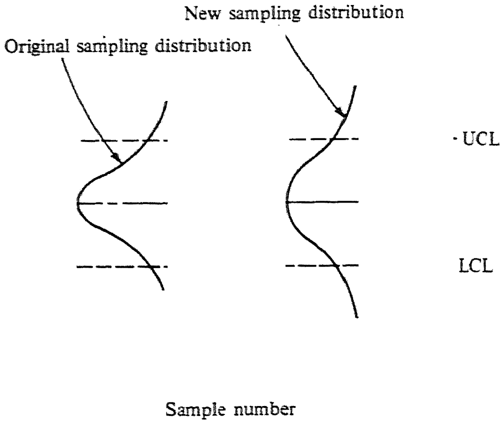


Fig. 16-10 Process variability.

the process mean has shifted, remedial action is taken. In addition to the likelihood of a shift in the mean, the process variability may also change.

Consider the control chart shown in Fig. 16-10. The upper control limits and lower control limits are established at $\pm 3\sigma_{\bar{x}}$ from the mean of the process. The original sampling distribution of the mean from which the control limits were set is shown. Suppose that the mean of the process does not change, but that its variability does. This new sampling distribution of the mean is also shown in Fig. 16-10. The chance that an out-of-control condition is signaled has increased. However, this time it is not the mean that has shifted, but the standard deviation.

As a result of this second source of variability, control charts are often kept on the standard deviation of the process. Each time a sample is taken, its mean is plotted on the $\bar{x}$ chart and its variability is plotted on a control chart for variability.

CONTROL CHARTS FOR ATTRIBUTES

Instead of taking an actual measurement on the characteristic being controlled, a simple good or bad, go or no-go, decision can be made. Control charts for these kinds of measurements are based on the binomial distribution, where the fraction defective in the sample is

$$\bar{p} = \frac{x}{n} = \frac{\text{number of defectives observed}}{\text{total number of observations}}$$

and $$= \sqrt{\frac{\bar{p}\bar{q}}{n}}$$

where S_p = standard error of the mean for sample proportions
$\bar{p}$ = fraction defective in the sample
$\bar{q}$ = fraction acceptable in the sample $(1 - \bar{p})$
n = sample size

The control chart is set up so that for, say, 99 percent control limits, the UCL (upper control limit) is $P + 2.58S_p$ and LCL (lower control limit) is $P - 2.58S_p$ where P is the population mean expressed as a fraction defective. In cases where the population mean is not known, the combined results of several samples can be used as a first approximation.

Case study: quality control in bank auditing Most loans granted by the Midtown National Bank require that collateral be pledged by the loan applicant. Frequently this collateral is in the form of securities, whose market values are continually changing. The job of the loan department is to ensure that at least a certain percentage of the loan is always covered. When the value of the collateral falls below this amount, the account is said to be *undermargined*. At that time, a call for additional collateral is sent to the customer.

Since it would be costly to insist that the loan department keep every loan properly margined at all times, a performance level of 5 percent has been set. As long as 5 percent or fewer of the outstanding loans are undermargined, it is assumed that the loan department is "under control."

Periodically the bank's auditing department must audit the loan department to verify that their accounts are properly margined and that the process is "in-control." The head of the auditing department, Norm Morris, has recently suggested that the bank employ statistical sampling procedures for these audits.

In a memo which he recently sent to those responsible for setting this policy, he explained that the full range of sampling strategies up to 100 percent inspection is open to them. He went on to say that if a 100 percent sample is taken, there would be no audit risk; they would, in fact, know precisely the percent of undermargined accounts. This strategy, however, does require a large staff; and as the bank grows, the staff will become even larger. Such a strategy is therefore quite costly. This cost would, in fact, be justifiable if the cost of operating above the 5 percent level were infinitely high. One would expect, however, that the higher the undermargined rate, the higher the expected losses from defaulted loans. But in no case would they be infinitely high. The risk incurred by sampling, then, may be worth it; the savings in audit time could be much greater than the expected costs of a higher than desired level of undermargined accounts. Essentially he was saying that the risk of accepting a process which actually has 7 percent undermargined accounts may be worth it. The precise nature of the risk will, of course, finally depend upon the values of AQL, α, LTPD, and β which are finally chosen.

Morris also told them that the magnitude of the risk could be tied to the size of the loan. For example, they might be inclined to take no risk for large commercial loans and a modest risk for consumer installment loans. To use statistical terminology, the audit department could stratify these loan accounts by dollar amounts and then determine the degree of sampling risk to which each stratum should be exposed. For example, three categories might include: (1) time loans of less than $10,000; (2) time loans greater than $10,000; and (3) demand loans. Since the risks that can be taken for each of these categories will be different, the sample sizes also will be different.

//// BAYESIAN QUALITY CONTROL

Up to this point we have taken the classical statistician's point of view. In specifying the criteria for the sampling plan (AQL, α, LTPD, and β), we presumably took into consideration the cost of making Type I and Type II errors, together with the cost of sampling. The selected plan somehow achieved the lowest sum of these costs. Nowhere, however, was it made explicit just how these costs were considered. In addition, any knowledge concerning the quality of the lot which the decision maker might possess prior to the sampling process went unused.

In the bayesian approach, these shortcomings are overcome. First, use is made of any prior knowledge. Consider a shipment which the purchasing agent of the G. E. Work Company has just received from one of his vendors. Based on past experience, he feels that there is a high likelihood that these parts are well within the desired quality limit. The bayesian statistician would insist that this knowledge, albeit intuitive, be used in the decision-making process. He would take a sample and use its result to revise his prior knowledge. Then he would either accept or reject the lot.

PRIOR PROBABILITIES

Let us look at how these prior beliefs can be expressed. Suppose that the full range of likely percent defectives for a particular process is 0, 1, 2, 3, 4, 5, and 6 percent. Furthermore, suppose that the decision maker feels that the likelihood of each of these occurring is .05, .10, .20, .30, .20, .10, and .05. These are shown in Table 16-3.

Table 16-3 Prior probability distribution

True process defective in percent p_i	*Prior probability* $P(p_i)$
0	.05
1	.10
2	.20
3	.30
4	.20
5	.10
6	.05

LOSS FUNCTION

The identification of the loss associated with sampling errors is the second shortcoming that bayesian analysis overcomes. Rather than specifying an AQL level and an LTPD level, the bayesian approach makes explicit the losses that can be incurred. We measure these losses as opportunity losses, where an opportunity loss is the difference between the cost actually incurred because of the decision and the cost which would have been incurred had the best possible decision been made.

Consider the following example. A manufacturer orders components in 2,000-unit lots. Whenever defective parts are uncovered, they are reworked at a cost of \$1 each. If the lot is rejected, a fixed cost of \$60 is incurred to cover clerical and handling costs.

If a lot is accepted, the costs of acceptance plus any necessary rework are

$$\$1 \times p \times 2{,}000 = 2{,}000p$$

If, on the other hand, the lot is rejected, the cost is \$60 regardless of the number of actual defectives. The point at which the decision maker is indifferent as to whether he accepts or rejects occurs when these costs are equal, that is,

$$2{,}000p = 60$$

or when $p = 3$ percent.

If we reject the lot when the actual percent defective is greater than 3 percent, the best decision has been made and no opportunity loss has been suffered. If, however, the lot is rejected when the percent defective is less than 3 percent, a needless \$60 cost is incurred when in fact we should have incurred only a cost of \$2,000p. This is a Type I error. Summarizing the opportunity loss associated with rejecting the lot, we have

$$\begin{aligned} \text{Type I error:} \quad & L_1(p) = 60 - 2{,}000p && \text{for } p \leq 0.03 \\ & L_1(p) = 0 && \text{for } p > 0.03 \end{aligned}$$

Now we turn to the Type II error. If we accept a lot which had an actual percent defective in excess of 3 percent, a cost of 2,000p is incurred when it was necessary to incur one of only \$60. If, however, we accept a lot where the actual percent defective is less than 3 percent, the opportunity cost is zero. We can summarize the opportunity loss of accepting the lot in the following way.

$$\begin{aligned} \text{Type II error:} \quad & L_2(p) = 0 && \text{when } p \leq 0.03 \\ & L_2(p) = 2{,}000p - 60 && \text{when } p \geq 0.03 \end{aligned}$$

We have now established a conditional loss function. *When* a good lot is rejected, $L_1(p)$ is the opportunity loss; and *when* a bad lot is accepted, $L_2(p)$ is the opportunity loss.

ERROR

Suppose that it has been decided to take a sample of 100 pieces (we will return shortly to a method for finding the optimal sample size) and reject the lot if the number of defectives in the sample is greater than c^*. The problem is to determine the critical value c^* which minimizes opportunity loss for that particular sample size.

Let us start by computing the total expected loss for the strategy $n = 100$, $c = 2$. First we find the probability of making an error for each possible percent defective. For example, if the lot percent defective is 0 percent, the possibility of making an error is zero, since it would be impossible to draw three or more defectives from the lot. If the lot percent defective is .01, a Type I error could be made; that is, a perfectly good lot may be rejected. We can compute this possibility from a binomial table. From the table in Appendix B we see that when a sample of 100 is taken and the true fraction defective is .01, the likelihood of three or more rejects in this sample can be computed by subtracting the probability that 0, 1, or 2 defectives will occur from one. In this case we have $1 - [.3660 + .3697 + .1849] = .0794$. This and the probability of also making a Type I error when the actual percent defective is 2 or 3 are shown in Table 16-4. If the true percent defective is either 4, 5, or 6, a Type II error can be made.

The probability of accepting a lot with 4 percent defective can be computed by summing the probability of 0, 1, or 2 defectives in the sample. From the binomial table we have $.0169 + .0703 + .1450 = .2322$. The probability of accepting a lot with 5 and 6 percent defective is computed in the same way and recorded in Table 16-4.

Now we can compute the total expected loss for this strategy. First the conditional loss is computed for each true fraction defective from the conditional

Table 16-4 Probabilities of errors for the decision rule $n = 100$, $c = 2$

True percent defective		*Probability that an error will be made*	
0	Type I error	.0	
1		.0794	probability of rejecting a good lot
2		.3233	
3		.5801	
4	Type II error	.2322	probability of accepting a bad lot
5		.1183	
6		.0566	

loss function. For example, if the true fraction defective is .01, the conditional opportunity loss associated with this Type I error is

$$L_1(p) = 60 - 2{,}000(.01)$$
$$L_1(p) = 40$$

In a similar fashion, the conditional opportunity loss associated with each of the other true fraction defective is computed and shown in Table 16-5. Next the conditional *expected* loss for each fraction defective can be computed.

$$\begin{array}{c}\text{Conditional}\\ \text{expected}\\ \text{loss}\end{array} = \begin{pmatrix}\text{conditional}\\ \text{loss}\end{pmatrix} \quad \begin{array}{c}\text{probability}\\ \text{that an}\\ \text{error will}\\ \text{be made}\end{array} \times \begin{pmatrix}\text{prior}\\ \text{probability}\end{pmatrix}$$

Finally, the total expected loss can be determined by summing the conditional expected losses. We can therefore conclude that if a sample of size $n = 100$ is taken and lots are rejected when $c \geq 3$, an expected cost of \$3.1822 will be incurred.

We are not through yet, because it might be that for a sample of size $n = 100$ another value of c might have associated with it a lower total expected loss. Therefore we must determine this loss for values of $c = 0, 1, 2, 3, 4, 5, 6,$ and so on. The probability that an error will be made is computed from the binomial table for each of these values of c and recorded in Table 16-6. Then the total expected loss is determined for each of these strategies just as was done in Table 16-5. The summary of this analysis is shown in Table 16-7. From this it can be seen that if a sample of $n = 100$ were taken, expected losses would be minimized when $c = 2$.

Thus far, what have we determined? That *if* a sample of $n = 100$ were taken, expected losses would be minimized if $c = 2$. But what about other

Table 16-5 Computation of expected loss when $n = 100$, $c = 2$

(1) *True fraction defective* p_i	*(2)* *Conditioned loss* $L(p)$	*(3)* *Probability that an error will be made*	*(4)* *Prior probability* p_i	*(5)* *Conditional expected loss*
0	\$60	0	.05	.0000
.01	40	.0794	.10	.3176
.02	20	.3233	.20	1.2932
.03	0	.5801	.30	.0000
.04	20	.2322	.20	.9284
.05	40	.1183	.10	.4732
.06	60	.0566	.05	.1698
			Total expected loss	3.1822

Table 16-6 Probabilities of errors for various decision rules with $n = 100$

Fraction defective	*$c = 0$*	*$c = 1$*	*$c = 2$*	*$c = 3$*	*$c = 4$*	*$c = 5$*
0	0	0	0	0	0	0
.01	.6340	.2642	.0794	.0184	.0034	.0005
.02	.8674	.5967	.3233	.1410	.0508	.0115
.03	.9524	.8054	.5801	.3528	.1821	.0808
.04	.0169	.0872	.2322	.4295	.6289	.7884
.05	.0059	.0371	.1183	.2578	.4360	.6160
.06	.0021	.0152	.0566	.1430	.2768	.4407

Table 16-7 Expected losses for different values of c with $n = 100$

$c = 0$	6.1031
$c = 1$	3.9864
$c = 2$	3.1822
$c = 3$	3.8158
$c = 4$	5.3068
$c = 5$	6.9877

sample sizes? Might a sample size of 50 or 200 be even less costly? Unfortunately there is no other way to find out than to go through this same analysis, starting with a sample size of 1 and continuing until the overall minimum is reached. Since larger samples are more expensive to take than smaller ones, sampling costs should be included along with the total expected loss for each sample size. As the sample size gets larger, sampling costs will increase while expected losses will decrease. For some sample size n^*, the sum of these costs will be minimized. We can therefore conclude that the optimal sample size is that sample size which minimizes the sum of sampling costs and expected losses.

It should be obvious that to find the optimal sample size is not a trivial chore; a lot of arithmetic is involved. If, on the other hand, a computer routine is available, the determination of the sample size can be a simple matter.

POSTERIOR ANALYSIS

The analytical framework developed thus far is used to determine the appropriate sample size. Once this is accomplished, a sample is taken and a sample result is observed. In bayesian analysis this sample result is then used to revise the set of prior probabilities. It is from this revised set of probabilities that a decision is made to reject or accept the lot.

Suppose that in our sample of $n = 100$ pieces exactly six defective items were found. With this new information we now compute a revised likelihood

Table 16-8 Revision of prior probabilities

True fraction defective p_i	*Prior probability* $P(p_i)$	*Conditional probability* $P(d\|p_i)$	*Joint probability* $P(d\|p_i) \times P(p_i)$	*Posterior probability* $P(p_i\|d)$
0	.05	0	0	0
.01	.10	.0005	.0000	.0000
.02	.20	.0114	.0023	.0373
.03	.30	.0496	.0149	.2422
.04	.20	.1052	.0210	.3414
.05	.10	.1500	.0150	.2439
.06	.05	.1657	.0083	.1349
			.0615	.9997

for each possible true fraction defective p_i. Said another way, we will revise our prior probabilities $P(p_i)$ to a posterior probability. This posterior probability can be written as $P(p_i|d)$, where d implies the occurrence of a sample result with d defective items. This revision can be accomplished by employing Bayes' rule.

$$P(p_i|d) = \frac{P(d|p_i)P(p_i)}{\sum\limits_{\text{all } i} P(d|p_i)P(p_i)}$$

The sequence of steps for the example $n = 100$, $d = 6$ is shown in Table 16-8. The conditional probability $P(d|p_i)$ is the probability of observing exactly d defectives, given that the true fraction defective is p_i. For example, the probability of exactly six defectives, given $p_i = .02$, is found to be .0114 from the binomial table.

Once the posterior probabilities have been computed, the total expected loss of rejecting and accepting can be determined. The expected loss associated

Table 16-9 Computation of expected losses after sampling has occurred

		Rejection		*Acceptance*	
p	*Posterior probability* $P(p_i\|d)$	*Conditional loss* $L_1(p_i)$	*Expected loss* $L_1(p_i) \times P(p_i\|d)$	*Conditional loss* $L_2(p_i)$	*Expected loss* $L_2(p_i) \times P(p_i\|d)$
0	0	60	0	0	0
.01	.0000	40	.0	0	0
.02	.0373	20	.7460	0	0
.03	.2422	0	.0	0	0
.04	.3414	0	.0	20	6.828
.05	.2439	0	.0	40	9.756
.06	.1349	0	.0	60	8.094
			.7460		24.678

with rejecting the lot is

$$\sum_{\text{all } i} L_1(p_i) \times P(p_i|d)$$

The expected loss associated with accepting the lot, on the other hand, is

$$\sum_{\text{all } i} L_2(p_i) \times P(p_i|d)$$

Whichever is lower becomes the chosen strategy. The computations for this problem are illustrated in Table 16-9. It can be seen that the lowest expected costs will be incurred if the lot is rejected. ////

SUMMARY

In the control of quality, there are very few occasions when it is necessary to completely inspect every item in the lot. Consequently a sampling plan is employed which reflects the decision maker's attitude about the risk which he is willing to incur. As the sample size is increased, the cost of sampling increases but the risk of passing defectives through the system decreases. The economic consequence of passing fewer defectives through the system is the savings of unnecessary work performed on defective parts and of the costs associated with defective parts reaching consumers. If the sample size, on the other hand, is decreased, sampling costs decrease but the cost of passing defectives increases. Difficult as it is, the designer must have the sum of these costs in mind when setting AQL, α, LTPD, and β. Ideally he would like to set these so that the sum of sampling costs and the cost of passing defectives through the system is minimized. From a practical point of view, as long as he gets close, he has done the job well.

QUESTIONS

Q16-1. How is good quality defined?

Q16-2. Differentiate between product and process quality control.

Q16-3. What factors influence the location of inspection stations?

Q16-4. What is sampling risk?

Q16-5. What factors influence the size of the sample?

Q16-6. How is a sampling plan related to AQL, α, LTPD, and β?

Q16-7. How is the outgoing quality of a sampling plan related to its incoming quality?

Q16-8. Why can sequential sampling be more economical than conventional sampling methods?

Q16-9. Compare attribute sampling and variable sampling. When is each appropriate?

Q16-10. When using control charts to monitor a process, two types of errors can be made. What are they?

Q16-11. Why is it necessary to monitor not only the process mean but also the process variability?

PROBLEMS

P16-1. Draw the OC curve for the following sampling plan.

n	100
c	6

P16-2. Find the sampling plan that will best meet the following objectives.

AQL = 3 percent
α = 95 percent
LTPD = 12 percent
β = 10 percent

P16-3. A manufacturer would like to specify a sampling plan which would have a high likelihood of accepting good-quality lots. To be more specific, he feels that lots of 2 percent actual defectives should be accepted by the plan 95 percent of the time. On the other hand, lots of poor quality, 8 percent actual defective, should be accepted only 10 percent of the time. Find a sampling plan which will achieve these objectives.

P16-4. Find the sampling plan which will best meet the following objectives.

AQL = 2 percent
α = 95 percent
LTPD = 8 percent
β = 15 percent

(*a*) Let α float.
(*b*) Let β float.

P16-5. Suppose that a very large lot whose incoming quality is 4 percent defective is processed through a sampling plan where $n = 100$ and $c = 4$. If lots of this quality are repeatedly processed through this plan, what will be the *average* outgoing quality? Assume that rejected lots are sampled 100 percent and only good items are accepted.

P16-6. Records indicate that the mean of a process has been 2.00 inches and its standard deviation .40 inch. Design a control system for this process such that the chance of making a Type I error for a sample size of 16 items is less than 1 percent.

P16-7. Establish a control system for a process which over the past has exhibited an average percent defective of 6. Set the sample size at 25 and the control limits such that the probability of making a Type I error is 5 percent or less.

P16-8. Establish a control chart for process variability when the average standard deviation of the process for samples of size 16 is $X_s = 5.0$ with its own standard deviation of 1.0. Set up control limits such that 95 percent of the time standard deviations are observed within these limits when the process is actually in-control.

P16-9. A manufacturer orders parts in lots of 1,000. Whenever defective parts are uncovered, they are reworked at a cost of \$2 each. If the lot is rejected, a fixed cost of \$80 is incurred.

(*a*) Determine the loss functions.
(*b*) Compute the total expected loss for the strategy $n = 100$, $c = 4$.
(*c*) If $n = 100$, what is the optimal value of c?

(*d*) Suppose that in the sample of $n = 100$ items exactly five defectives were uncovered. What is the optimal strategy? In reaching your decision, utilize the decision maker's prior estimates of process defective.

True process defective in percent p_i	*Prior probability* $P(p_i)$
0	.02
1	.05
2	.20
3	.46
4	.20
5	.05
6	.02

BIBLIOGRAPHY

Bowman, Edward H., and Robert Fetter: *Analysis for Production and Operations Management*, 3d ed., Richard D. Irwin, Inc., Homewood, Ill., 1967, chaps. 5 and 6.

Buffa, Elwood S.: *Models for Production and Operations Management*, John Wiley & Sons, Inc., New York, chaps. 6, 7, and 8.

Dodge, Harold F., and Harry G. Romig: *Sampling Inspection Tables—Single and Double Sampling*, John Wiley & Sons, Inc., New York, 1944.

Duncan, Acheson J.: *Quality Control and Industrial Statistics*, 3d ed., Richard D. Irwin, Inc., Homewood, Ill., 1965.

Sasaki, Kyohei: *Statistics for Modern Business Decision Making*, Wadsworth Publishing Company, Belmont, Calif., 1968, chaps. 11, 12, and 13.

17

Management Information Systems

INTRODUCTION

Before the availability of computers, all records, including accounts receivable, payables, inventory, payroll, personnel, and the billing of customers, were kept by bookkeepers and clerks. Record-keeping systems were frequently quite sizable, inefficient, and subject to human error.

In the early years of computers it was, therefore, not surprising to find the focus on these applications. The potential payoff seemed promising. The results, in fact, bore this out. It was proved that these computer methods held substantial advantages over the old labor-intensive methods. Among them were speed, accuracy, and cost savings.

In these applications no decision making is involved; there is no choice between alternatives. At most, some arithmetic operations are performed and the output summarized on a record or stored in the computer's storage system. For example, in preparing payroll the employee's identification number, hourly wage rate, number of dependents, weekly retirement deductions, union dues, and so on, are recorded and stored on tape. At the end of the week, his regular and overtime record is submitted and the computer, upon reaching his file, computes his pay, makes the necessary deductions, and supplies him with the appropriate records and paycheck. For this and the other early applications of the computer, therefore, only simple data processing was involved.

As the computer assumed this information-processing role, it became apparent that the timely source of data which it generated could possibly be useful for decision-making purposes. Data-processing departments then started to supply management with voluminous streams of data. There are several problems, however, with this approach. The first is probably one of the most important: too much data is frequently as bad as none at all. A busy manager is not likely to pour through a several-hundred-page daily load status report before making his daily schedule. Secondly, data in this form are not necessarily the kind of information that is required for decision-making purposes. Both of these problems point to the fact that the use of the computer in a management information system capacity requires at least the coordination of the system with the potential decision makers and perhaps even the collection of data other than those which are collected for the accounting and financial reporting system.

These problems led to a fresh inquiry into the way in which information should become involved in the decision-making process. Out of this came the discipline called *management information systems.* Its focus ranges from simple applications, where relevant financial and accounting data are effectively summarized and presented, to systems where complex decision problems are solved by computer routines.

The ultimate, one might suggest, would be to have all decisions programmed so that they could be made automatically. In this way complex interactions between decisions could be incorporated in the decision-making models to ensure global optimization of the firm.

Such a system is at least a long way off—but, more probably, unattainable. The reason for this is that the environment within which decisions are made is not only continually changing but at any one time may be too complex to model in its entirety. Quite attainable, moreover, is a management information system lying somewhere short of this extreme: a system which is not totally automatic and one which requires the decision maker to take an active role.

The more modest management information system is the focus of this chapter. In the following sections we will consider what a management information system is and what it is not. Then we will briefly examine some general characteristics of computer systems that act as the heart of the MIS. Finally we will take the major decision areas in operations management and develop a management information system centering on their related problems.

WHAT IS A MANAGEMENT INFORMATION SYSTEM AND WHAT IS IT NOT?

DEFINITION OF MIS

A workable definition, and one that will set the stage for the rest of this chapter, is the following: "An MIS is an automated system which presents information, both internal and external to the business, that aids in making a specific set of routine decisions."[1]

There are two aspects of this definition that warrant closer scrutiny. First, the purpose of the MIS is to aid decision making and not to automate the decision-making process itself. As a result, the decision maker is seen as having a dominant role, not the MIS. This in turn means that any attempt to design a system without the support and involvement of management is most often pointless. The second aspect of the definition is that the MIS should focus on routine decisions only. What this implies is that the determination of what should be included in an MIS is an economic decision by itself. The inclusion of any set of decisions, be they inventory, capital budgeting, scheduling, or line balancing, must be justified on a benefit/cost basis. There is a cost of including them in the system which consists of the model-building cost, then the cost of collecting and maintaining the relevant data base, and finally, the computer costs. The benefits, on the other hand, include the stream of savings that the new system will accrue over the old. Only those decisions whose benefit/cost ratio is attractive should be included.

The potential for MIS in an operational system is substantial. Generally the kinds of decisions that are made during the design, operation, and control of a transformation process exhibit much more structure and are much more routine than decisions made elsewhere in the firm. For example, there would be less controversy over the conceptualization, solution, and implementation of a model designed to solve an inventory problem than of one for setting a marketing strategy. Consequently much of the effort in MIS has been directed at operational problems. There have, indeed, been countless successful applications. The record, however, has not been unblemished.

PITFALLS IN DESIGNING AN MIS

Experience has uncovered several erroneous assumptions which help account for the failure of countless MIS projects.[2] Reviewing them should provide some useful insight into effective MIS design.

The first wrong assumption is that *managers suffer from a lack of relevant information*. There is some truth to this; however, it seems more realistic to say that they suffer more from an overabundance of irrelevant information. Therefore, the focus of an effective MIS should not be on supplying relevant information but on eliminating irrelevant information. If the emphasis was on relevant information, the information specialists would busy themselves with generation, storage, and retrieval of information. The consequence of this would be a profusion of data, difficult to absorb and analyze.

The second erroneous assumption is that *MIS systems should be based on the kinds of information that management needs*. Since most managers do not understand the structure displayed in some decision situations, they will likely ask for more information than is needed. In this way they play it safe. An excellent example of this is given by Ackoff. He tells the story of a group of marketing managers who were asked for the variables which they thought were relevant in explaining the sales volumes of filling stations. Almost 70 variables were identified. The market research team subsequently built a regression model in which 35 variables were found to be statistically significant. Later an OR team built a simple regression model with traffic flow as the only independent variable. It performed best of all.

---

[1] Harry Stern, "Information Systems in Management Science," *Management Science*, vol. 17, no. 2, October 1970, p. B-121.

[2] Taken from Russell L. Ackoff, "Management Misinformation Systems," *Management Science*, vol. 14, no. 4, December 1967.

The point is that any analysis should start with the decision process. What are the kinds of decisions made? What information is required? Once these questions are answered, only those data which are needed should be supplied. The management information system should be a subsystem of the decision process and not the other way around.

The third erroneous assumption is that *if a manager has the information he needs, his decision making will improve.* That this may not be true can be seen from a simple example. Consider the traveling salesman problem, where a salesman must be routed to each of 10 cities. If a manager is given the driving time between each possible pair of cities, it is very unlikely that he could structure a tour such that driving time was minimized. With over 3,000,000 possible tours, a more effective approach would be to either provide him with a decision rule or, by using an appropriate model, supply him with the optimal tour.

Therefore, whenever information is provided, it is necessary to consider how well these decision makers can use these data. If the decision process is structured and complex, the use of models and decision rules should be carefully considered.

Finally we turn to the erroneous assumption that *a manager does not have to understand how the information system works to use it.* If, in fact, the decision maker does not understand how the information system works, he probably had no role in its design and currently has no role in its control. As a result, few decision makers would place much confidence in its use. If, on the other hand, he were to use the output, the decision maker would be in the position of being controlled by the information system rather than controlling it.

Understanding the MIS is not easy. It takes some knowledge of management science and the computer—precisely the kind of background you are gaining by reading this book. Most important, it takes confidence to ask the information system people the right kinds of questions. What are the assumptions of the model? How, where, and when are the data collected? In addition, a continual feedback process is essential so that the information system people may be kept sensitive to the dynamic nature of the decision environment; there is nothing quite so irrelevant as a decision model which was designed two years ago to solve a problem which has dramatically changed over time. Stagnant information systems such as this are doomed to failure. The only way to avoid this pitfall is to encourage understanding and participation by information specialists, management scientists, and managers.

Now that we have considered the role of the decision environment on the one hand and that of information collection and presentation on the other, we turn to the link between them: the computer systems. To do this subject justice takes much more time and space than this chapter can allow. Only a few very basic and general concepts will be covered.

CHARACTERISTICS OF COMPUTER SYSTEMS

A computer system has three basic units. These are its input, output, and central processor. The input unit is the means by which data are supplied to the computer (see Fig. 17-1). Input devices include punched-card readers, magnetic tapes, and disks. The central processing unit (CPU) is usually considered the heart of the computer and is composed of a control, arithmetic, and memory unit.

When a set of operations is to be performed by the computer, a set of instructions called a *program* is entered through the input unit and stored in memory. In addition, any data that are required are also entered and stored. To be acted upon, the data and program are moved into the control unit and the appropriate arithmetic and logical operations are executed.

Finally we have the computer output units. Output can be presented in several ways, including punched card, tape, disk, printed records, and video display.

Two distinguishing features between computers include their storage capacity and their response time. Storage capacity is the ability of the computer to both temporarily and permanently store data. For example, in an inventory control system, the computer must be capable of storing inventory levels, data for forecasting purposes, reorder levels, and reorder quantities. When hundreds of thousands of items are carried in inventory, a sizable storage capacity is required. The second distinguishing feature is the response time. This is the time between when the input is supplied to the computer and when the output becomes available.

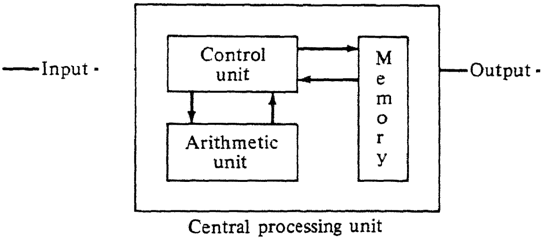


Fig. 17-1 Central processing unit.

BATCH PROCESSING

The predominant system in the past has been batch processing. In this system, the computer is totally dedicated to the job being run. That is, other jobs must wait in a queue until the one being processed is completed. This approach has two major drawbacks, including inefficiency of the central processor and turnaround time.

In a moderate-sized batch facility, a program and data are entered into the computer's memory, the program is executed, and the output is eventually supplied. During much of this process, the CPU is entirely dedicated to one job, and as a result a considerable amount of idle time is incurred. For example, if the program requires that data which are stored on a disk file be processed, the CPU sits idly by as the search for the data is undertaken. In strict batch processing, little can be done to overcome this problem.

The turnaround problem is the second major drawback of batch processing. Usually a job is brought to the central computing facility and output is picked up some minutes or hours later. The wait depends upon the number of jobs in the queue and on their length. It is not unusual for this process to take 24 hours. This delay, together with the inconvenience of dropping off and picking up the job, casts an unfavorable shadow over batch processing from a user's point of view.

For many applications in a management information system, the computer must be easy to use, convenient, and quick to respond. A batch system does not fulfill these requirements. There are, however, some kinds of information which are needed at infrequent intervals which can be processed by a batch system. These include monthly financial statements, monthly inventory summaries, and so on.

REMOTE BATCH

One way in which the turnaround time for batch systems has been reduced is by the use of remote batch processing. Remote terminals are placed in convenient locations and thereby provide remote input-output access to the computer. Jobs are entered from them and are transmitted to the main computer facility, where they take their place in a queue and wait their turn. Other than the convenience of remote terminals, the system basically operates as a batch process.

MULTIPROGRAMMING

Multiprogramming is a technique which allows the CPU to operate on more than one program, thereby overcoming a basic inefficiency of the batch process. It works in the following way. If the central processor is idle at any time during the execution of a program, the program that is next in queue is worked on temporarily. When the first program once again requires the use of the CPU, the processor switches from the second program to the first. For example, if the primary program directed the computer to search for data in a storage file, during the time it would take to seek and retrieve this information, the CPU would switch to another program. A major advantage of multiprogramming, therefore, is that the CPU is kept busy most of the time. To use such a system, however, is still inconvenient. Although jobs in the queue may be partially completed, the first one receives primary attention until it is completed, then the next one moves to the primary position and it is processed until complete, and so on. A three-minute job with 25 jobs ahead of it may still have to wait 24 hours for completion.

ENQUIRY SYSTEMS

An enquiry system operates in real time. That is, the remote console operator receives the impression that the computer is responding nearly immediately to his inquiries and that he is the sole user. Typical of the use to which these systems are put are stock quotation services and airline reservation systems. In these applications a trivial amount of computation but a considerable amount of search for particular items stored in memory is required. Many of these systems are multiprogrammed, since a substantial portion of system time is spent searching for data.

GENERAL-PURPOSE TIME-SHARE SYSTEMS

In a general-purpose time-share system, the user has the impression that he is the only one using the machine and that the responses to his inputs are nearly instantaneous. This is accomplished by giving each user a slice of time. In this time slice he has complete use of the CPU, after which the next user has exclusive use, and so on, the CPU eventually returning to the first user. The cycle happens so fast that the user gets the impression of continuous, exclusive use.

The system is general-purpose in that the user can retrieve information, use available computer codes, or write his own programs in BASIC, COBOL, FORTRAN, and so on. These characteristics are especially important to the design of management information systems.

In addition to being general-purpose, the turnaround problem is solved. With a time-share system, for example, the consequences of making a change in a shop schedule can be determined almost instantaneously.

Time-sharing is especially attractive for interactive decisions. This is the situation where the decision maker requires some time to think, makes an entry into the remote terminal, receives a response, devotes more time to thinking, makes another entry, and so on. This alternating sequence is a useful concept in the design of an MIS.

EXAMPLES OF MIS

Most of the experience in developing management information systems has been limited to well-defined groups of problems. An example of this is the Sabre reservation system which is operated by American Airlines. This enquiry system is used to manage flight reservations. At the center of this system are some 1,200 terminals located in reservation offices throughout the country. They are in turn linked to a central computing facility through the telephone network. When an enquiry is made for seat availability on a particular flight, instantaneous up-to-date information can be obtained. If the customer desires the seat, his passage can be assured and the load status revised.

Without such an information system, the ability to verify seats would take much longer. Therefore, the timeliness of the data is one justification for the expenditure of some $20 million; customers must be given a confirmation while they are on the phone. A second justification is that the inventory is perishable. If a seat should become available as the result of a cancellation, its immediate entry into the load status of the plane might lead to its resale. In a less efficient system, it might go unsold.

An extension of this system is currently being considered. Into the remote-entry device would be placed the customer's desired date and time of departure, together with his destination. The computer would then determine the best available schedule which most nearly meets his departure objective.

Another example of a specialized MIS is found in banking. Here, remote units are located at tellers' cages. Bank transactions which are carried out at the window are transmitted to a central computer. In this way several customers can be served from several different locations and the file on each customer instantaneously updated.

In judicial systems, MIS is being used for court scheduling, while the police use it for identifying and finding criminals. Its use in investment analysis has been growing rapidly. Amstutz[3] reports on an investment firm that has been using it for some eight years. He points out that even in investment analysis implicit management theories can be discarded in favor of explicit concepts.

Applications of MIS to health maintenance and diagnosis seem inevitable, for in this field is found a substantial amount of data collection, processing, and decision making. The functional components of this process involve the following five stages:[4]

1. The acquisition of data to establish health status
2. The comparison of these data with stored knowledge
3. The decision as to:
    a. What added data are needed?
    b. What treatment is indicated?
4. The administration of the treatment and the acquisition of follow-up data
5. The addition of all of this experience to stored knowledge

---

[3] Arnold E. Amstutz, "Information Systems in Management Science," *Management Science*, vol. 15, no. 2, October 1968, p. B-99.

[4] Harry Stern, "Information Systems in Management Science," *Management Science*, vol. 16, no. 10, June 1970, p. B685.

The application of MIS to these stages seems promising. Given the urgency of health care and our changing set of national priorities, much progress in this direction can be expected over the next few years.

Countless other examples can be found in the factory setting. The most prevalent are associated with forecasting, inventory, and scheduling problems. The scheduling system illustrated in Chapter 15 represents one of the more ambitious undertakings.

The systems which we have discussed so far are partial information systems. In the next section, an MIS for an entire operational system will be generated. Although none quite so pervasive is being used by industry, the intent of the section is to suggest what can be done and how these subsystems can be integrated into a functioning whole.

INTEGRATED MIS

For many the concept of a full-blown management information system is vague without the benefit of a supporting illustration. In this section, a management information system will be designed for a large manufacturing company producing a variety of products, both to order and for inventory. The focus will be on the kinds of decisions that *should* be made, on the information required to make these decisions, and on the models which might be helpful during the decision-making process.

This section will also serve to integrate the topics covered in previous chapters. Therefore you should begin to get the feel of how those previously isolated topic areas can be integrated into a management system—how, for example, sales forecasting, aggregate scheduling, shop loading, dispatching, and inventory control are related through a complex web of information flows.

To consider the entire task from long-range planning to shop control at once would be overwhelming. Instead we develop the system in three stages and then combine those stages into one management information system. The artificial classification which separates these stages is the planning horizon. The first stage centers around the decisions necessary in long-range planning. Here the planning horizon is one year and beyond. The second stage focuses on aggregate planning, where the planning horizon is from one month to one year. Finally, the third stage is concerned with the short term, a period usually considered to be less than a month.

LONG-RANGE PLANNING

Forecasting The essential ingredient in any plan is a forecast for the future level of activity. In long-range planning, the purpose of this forecast is to identify the kinds of products or services that will be demanded as well as some indication of volume. Unfortunately the traditional forecasting techniques, such as exponential smoothing and econometric analysis, can be of very limited help. This is so because long-range planning is concerned with the nature of the marketplace several years away, and in this length of time the whole structure of demand may change. Therefore it is probably of little value to use exponential smoothing techniques, which extrapolate historical time series data into the future, or econometric analysis, which bases its predictions on historical cause-and-effect relationships. Long-range forecasting methods will probably continue to rely, as they have in the past, on insight, intuition, experience, and an occasional educated guess. Consequently, there is very little that a management information system can do to help. For example, at best it can be used to summarize available competitors' data, industry data, and industry forecasts.

Process strategies Once these estimates have been made, they must be combined with preliminary engineering and manufacturing data to determine the alternative process strategies which can be employed to meet this type and level of output. This sequence of information flow is illustrated in Fig. 17-2.

Capital budgeting Since the alternative solutions to the production plan require a share of the scarce pool of capital, and since marketing and engineering are also making demands on these funds, some interaction is required when these funds are allocated. In Chapter 2 it was suggested that the common denominator by which all projects, including replacements, compete for scarce funds be the benefit/cost ratio. In this way the most profitable alternatives, regardless of functional identity, can be selected first, then the next most profitable, and so on, until the budget is exhausted.

The capital-budgeting process can very effectively be included as part of an MIS. Into remote terminals can be entered the relevant data. The computer could then execute the necessary calculations and present the output data perhaps as decreasing benefit/cost ratios or as internal rates of return. Perhaps the more complex models which employed linear programming and simulation to cope with constrained decision making and uncertainty would be even more appropriate.

Project scheduling After the capital-budgeting process is complete, a set of projects emerge which form the project plan for the future. To plan and control the execution of these projects from their design stage through the point in time when they become operational, some form of project scheduling should be employed. In Chapter 7, several project-scheduling techniques, including PERT and CPM, were presented. We learned that the choice of a particular technique depends upon the predominant factors in the actual situation.

The formulation of the project network and the subsequent control of the project can be included as part of an MIS. If status reports on each activity are periodically entered, critical paths can be recomputed and project control can therefore be based on the latest information.

Facilities location If the project requires the expansion or relocation of physical plant facilities, a facilities location study is usually initiated. In some well-

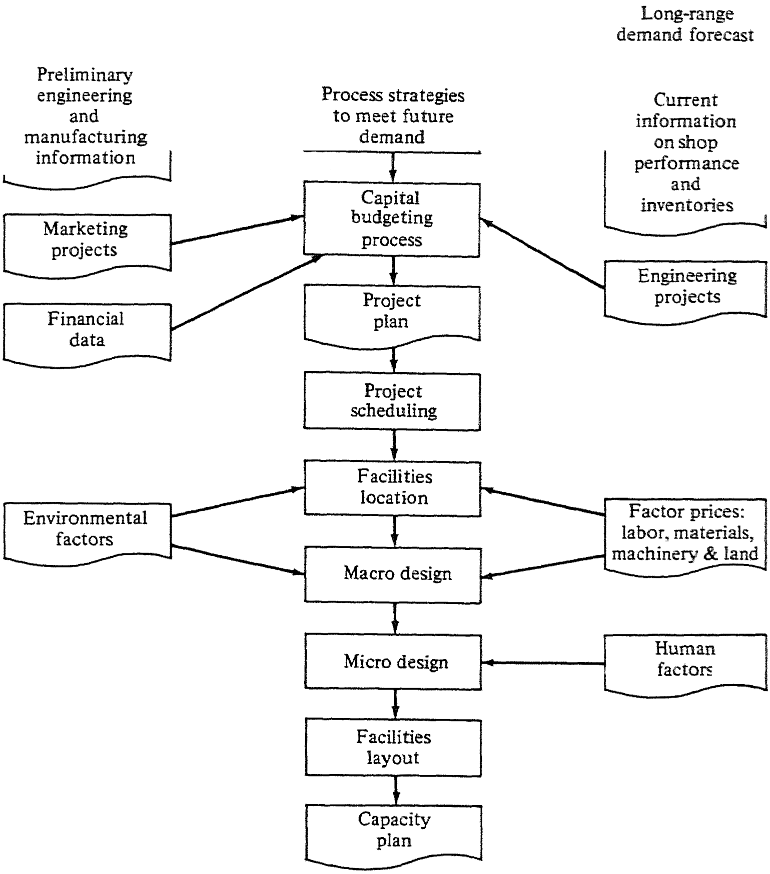

Long-range demand forecast
Preliminary engineering and manufacturing information
Process strategies to meet future demand
Current information on shop performance and inventories
Marketing projects
Capital budgeting process
Financial data
Engineering projects
Project plan
Project scheduling
Facilities location
Environmental factors
Factor prices: labor, materials, machinery & land
Macro design
Micro design
Human factors
Facilities layout
Capacity plan

Fig. 17-2 Long-range planning.

defined problems, a management information system can be helpful in arriving at the location decision. Examples of two such situations were given in Chapter 9. The first was a company where transportation and manufacturing costs dominated the decision, and the second was a telephone company where the location of a central switching office significantly affected the investment in copper wire between the switching office and subscriber telephones.

Macro design The next stage is macro design. This topic was covered in Chapter 5 and focused on such problems as the choice of either a product or a process layout, the design of the process, selection of equipment, major make or buy decisions, and material handling. As we discussed in that chapter, this is not a "once only" effort to be initiated for new projects. It should be structured as a periodic review system where new alternatives are evaluated in response to either new technological process developments or changes in the prices of labor, materials, or equipment.

Consider the following example of how a management information system can provide valuable assistance in solving macro design problems. In a job shop the feedback from the shop has shown a chronic bottleneck situation at several work stations. As a result jobs have not been meeting their due dates. One of the several suggested solutions to the problem is to purchase a $600,000 numerically controlled machine. It should be able to absorb many of the jobs which are now being processed on several different machines. If the relevant job arrival times and service times have been stored, simulations for this and other alternatives could be run to determine the effectiveness of each strategy.

Major make or buy decisions are usually considered in this phase. Since the price breaks, quality, and delivery of purchased items are continually changing between vendors, the firm should maintain some routine procedure for periodically ensuring the best strategy. Without it the tendency is to stay with the same vendors. This review process could also be included as part of the MIS.

Material handling problems of the kind illustrated in Chapter 5 also lend themselves to quantitative analysis. If they recur frequently enough, a materials handling model could be incorporated as part of the MIS.

Micro design After macro design, focus shifts to the micro aspects of the process. Involved here are the determination of production standards, work methods, and finally line balancing, which acts as the integrating element.

In Taylor's era of scientific management, the micro design of the production process was approached from a mechanistic point of view. There was very little consideration directed toward the human element in the process.

Today, however, in process design considerable emphasis is given to human factors. As a result, the design of the process is less structured than under the Taylor approach. Usually, however, the design is accomplished by an iterative process. For example, in line balancing the initial cycle time for each work station may be selected in response to considerations of job scope and depth. Then the line is balanced using some optimization technique or heuristic which could well be part of an MIS. The resulting composition of work at each station is scrutinized and perhaps another cycle time is chosen, followed by the line-balancing routine, followed by another cycle time, and so on.

Facilities layout A facilities layout study must be undertaken not only when new facilities are planned, but also periodically in response to a new product mix, new levels in demand, or changes in the process. Since the alternative strategies are usually numerous, a decision model incorporated as part of an MIS could be quite useful. Several heuristics have been successfully employed by industry, one of which was presented in Chapter 8.

Capacity plan The outcome from this long-term strategic planning phase is a capacity plan showing a suggested strategy to be used over the next five or perhaps even the next 10 years. The plan includes the chosen projects, increase in physical plant capacity, new plant locations, equipment purchases, replacements, and even major changes in work force levels and inventory strategy. Its detail, furthermore, is greatest for the immediate future and decreases as the plan unfolds into the long-range future.

AGGREGATE PLANNING

In the aggregate planning phase, production plans are made for the period between the next one and 12 months. If the firm's manufactured output is destined to stock finished goods inventory, the starting point for the plan is a sales forecast. If, however, the firm manufactures only upon receipt of an order and customers are willing to wait for their special orders, aggregate planning can be based on orders. Sometimes a combination of both situations can occur. These two possible inputs are shown in Fig. 17-3.

Forecasting At this stage an MIS can provide valuable forecasting assistance. Using either exponential smoothing or econometric forecasting techniques, demand forecasts can be made. In either case it will be necessary to maintain some data bank. When econometric forecasting is used, however, one can find commercial time-share companies which offer data banks of over 6,000 economic variables which can be used as predictive variables in the forecasting model. This is usually an attractive alternative compared with generating and maintaining one's own data bank for forecasting purposes.

Aggregate requirements Now the sales forecast and customer order information, engineering and manufacturing data, and stock status can be combined to determine the requirements over the planning horizon. Usually the requirements are stated in aggregate dimensions, such as gallons of paint, number of TV sets, number of court cases, number of surgical operations, and so on. The

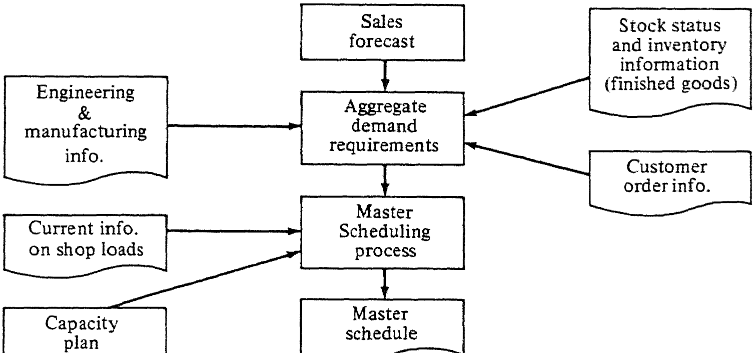

Sales forecast
Stock status and inventory information (finished goods)
Engineering & manufacturing info.
Aggregate demand requirements
Customer order info.
Current info. on shop loads
Master Scheduling process
Capacity plan
Master schedule

Fig. 17-3 Aggregate planning.

detail, which would include the different kinds of paints, the variety of TV sets, the kinds of court cases, or the type of surgery, is left for the next stage in the planning process.

There is usually good reason for this aggregate focus. Since this plan is drawn up some months before the orders materialize, it should be flexible enough to accommodate a variety in the product mix.

Master schedule Now the master schedule can be compiled. As inputs we include the aggregate demand requirements, the capacity plan, and the current information on shop loads. The capacity plan supplies the capacity levels over the planning horizon. If demand is undergoing a secular increase, capacity will probably be increased over time. As this capacity becomes available, it will be incorporated into the scheduling plan. Finally, information on current shop loads is supplied so that the initial conditions from which the schedule unfolds will at least reflect the current status of the shop.

The master schedule specifies when jobs will be initiated and completed, the production and work levels, make or buy decisions, undertime and overtime, hiring and firing, and inventory levels. It is therefore a production plan whose objective it is to meet a demand pattern as economically as possible.

Because of constraints such as capital, labor, physical capacity, and management, and because of fluctuating demand requirements, aggregate planning can be a complex problem. In Chapter 13 linear programming, transportation, and quadratic techniques for aggregate planning were illustrated. A model of this nature included as part of the MIS can, therefore, be quite useful.

SHORT-TERM SCHEDULING AND SEQUENCING

Now we are down to planning for the month in which production will take place. At this stage tremendous detail must enter the planning process; each part, subassembly, and final assembly for every job found on the master schedule must be scheduled. Without an MIS, this stage is frequently chaotic and inefficient.

Job explosion The first step in scheduling is to explode the jobs found on the master schedule. This explosion essentially decomposes the job into all of its basic parts and subassemblies. This is accomplished by merging the master schedule with the manufacturing information file. Also considered are changes in customer orders or special manufacturing instructions in response to custom orders. These relationships are shown in Fig. 17-4.

Requirements Once the total requirements are determined, a check is run on the quantity of these parts in inventory. When a sufficient quantity is available, the requirement is met from stock. When there is an insufficient quantity, an order is entered on the short-term schedule. It, however, must be scheduled in such a way that it will be ready to be combined with the other parts to ensure on-time completion of the job. You will recall from Chapter 13, whenever

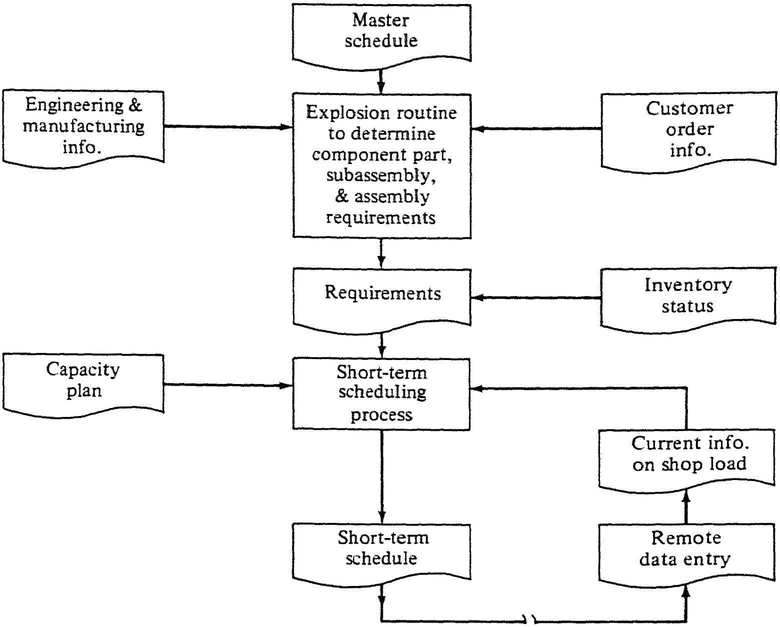

Master schedule
Engineering & manufacturing info.
Explosion routine to determine component part, subassembly, & assembly requirements
Customer order info.
Requirements
Inventory status
Capacity plan
Short-term scheduling process
Current info. on shop load
Short-term schedule
Remote data entry

Fig. 17-4 Short-term scheduling.

parts are manufactured within the firm, an EOQ supply is not necessarily ordered. The reason for this is that order quantities affect the shop load in addition to the usual inventory costs. Therefore these should be considered before an order quantity is entered. If the order for parts is placed outside the firm, the conventional inventory analysis of Chapter 14 applies.

Sequencing Once each part or subassembly has been assigned to a work center and its approximate starting time has been established, focus shifts to the work load at each of these work centers. The scheduling problem is now down to its finest detail; here we determine the order in which the jobs at each machine are to be processed. This topic was covered in Chapter 15, where models were presented for determining both a theoretically optimum solution and less-than-optimum solutions based on heuristic techniques.

Shop control When orders are released to the shop, the information system should automatically furnish the following:

1. Raw material and inventory requisitions
2. The necessary engineering and manufacturing drawings, either on microfilm or accessible by video display units with options for hard copy
3. Routing instructions
4. Job identification card

Since jobs seldom flow smoothly through the shop and meet their expected completion dates without incident, control measures must be instituted. To accomplish this, the progress of the job through the shop must be monitored. Whenever the job is started or completed at a machine center, the job card and location information are entered into remote data units placed throughout the shop. Therefore the whereabouts of every part or subassembly is known at all times. If the completion of the job becomes jeopardized, the appropriate action can be taken. This might consist of changing a part's priority.

A rather elaborate control system was presented in Chapter 15. The output of this control process included several reports which allowed for easier and more effective day-to-day scheduling, as well as more effective control over the system.

In Fig. 17-4, we can see where this feedback via remote data entry devices updates the current shop load and status file. It, in turn, is fed into the short-term scheduling process to generate a new, up-to-date schedule. As in the example presented in Chapter 15, this iterative process can take place daily. The day's manufacturing activities can therefore be planned from reliable data. Without such a system, the foreman typically responds to pressures from expediters and has little information, if any, on the change in priority of jobs which he neglects or pushes back in the schedule.

THE THREE STAGES COMBINED

In Fig. 17-5 the three stages in the planning process have been combined into a comprehensive management information system. Added to this system is the inventory planning and control stage. It illustrates another valuable feature of an MIS. By utilizing remote order entry units, which could be located in sales offices throughout the country, salesmen could make instantaneous price and delivery inquiries. For example, a customer who is interested in a particular product would like to know its probable delivery date. The salesman could gain access to the firm's central computer through his remote data entry device. The central computer can then determine whether the product is in stock or, if it is not, when the stock is scheduled to be replenished. If the customer is satisfied by this response and the sale is made, it can be entered into the remote unit and the product, either in stock or in-progress, can be assigned to that customer.

SOME MANAGEMENT IMPLICATIONS

Perhaps the role that the computer will play in management decision making will change the kind of leadership and management style that has been traditionally required. In the past, decision makers had to exhibit a skill for working with vague data. They then had to somehow reach a conclusion from these data, win approval of their plans, marshal the necessary resources, and initiate action.

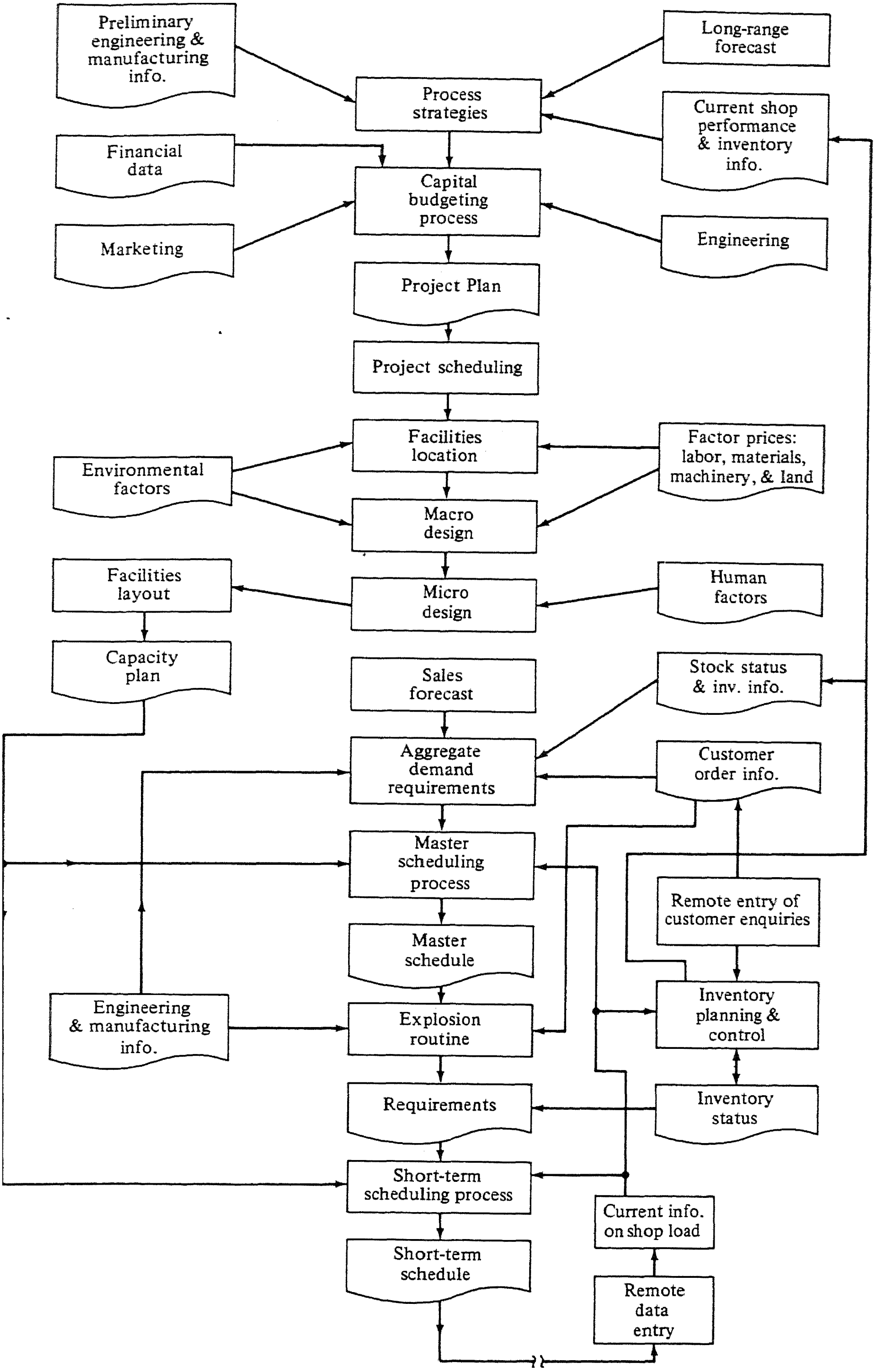

Preliminary engineering & manufacturing info.
Long-range forecast
Process strategies
Current shop performance & inventory info.
Financial data
Capital budgeting process
Marketing
Engineering
Project Plan
Project scheduling
Facilities location
Factor prices: labor, materials, machinery, & land
Environmental factors
Macro design
Facilities layout
Micro design
Human factors
Capacity plan
Sales forecast
Stock status & inv. info.
Aggregate demand requirements
Customer order info.
Master scheduling process
Remote entry of customer enquiries
Master schedule
Engineering & manufacturing info.
Explosion routine
Inventory planning & control
Requirements
Inventory status
Short-term scheduling process
Current info. on shop load
Short-term schedule
Remote data entry

Fig. 17-5 Major information flows.

Argyris[5] points out that an MIS could change this managerial style quite drastically. First, the new-style manager could experience a reduction in his freedom of movement. Both the formal and the informal aspects of the environment will tend to be much more explicit. Secondly, the manager could suffer from a feeling of psychological failure, since the MIS will be making many decisions for him. It sets his goals, it specifies the actions necessary to achieve these goals, and it evaluates his performance. Third, leadership in this new environment may no longer be as strongly related to the manager's political power within the organization as it is related to his competence. Fourth, the decision maker will no longer feel as essential to the management team as he has felt in the past. Since everything will be done in a planned or rational way, he will not feel as powerful as he did before.

Although Argyris is warning us about the consequence of a fully integrated and automated MIS, some of these caveats are still worth considering. Even in management information systems, which look at but a small subset of decisions, some of these behavioral problems are likely to emerge. It is important that management be sensitive to them.

We can conclude, then, that the success of an MIS will depend upon decision makers in two ways. First, it depends on their ability to work effectively with the information systems people and with the management scientists. Secondly, it depends on their ability to adapt to a new management style quite unlike anything we have known in the past.

QUESTIONS

Q17-1. Experimental emergency health care centers have been established at locations where none have previously existed. One such center was located at a large international airport.

In attendance at this location was a nurse who was in constant contact with the central hospital via a closed-circuit TV network.

What information flows and decisions would you consider in the design of this fairly sophisticated MIS?

Q17-2. How would you design an MIS for a supermarket chain?

Q17-3. Prepare an MIS proposal for a wholesale liquor dealer with three warehouses located in different cities.

Q17-4. How would you design, as part of a bank's MIS, an automated procedure for screening personal-loan applicants?

BIBLIOGRAPHY

Argyris, Chris: "Management Information Systems: The Challenge to Rationality and Emotionality," *Management Science*, vol. 17, no. 6, February 1971, p. B-275.

Betaque, Norman E., and G. A. Gorry: "Automating Judgmental Decision Making

---

[5] Chris Argyris, "Management Information Systems: The Challenge to Rationality and Emotionality," *Management Science*, vol. 17, no. 6, February 1971, p. B-275.

for a Serious Medical Problem," *Management Science,* vol. 17, no. 8, April 1971, p. B-421.

Davis, Gordon B: *Computer Data Processing,* McGraw-Hill Book Company, New York, 1969.

Dearden, John, F. W. McFarlan, and W. M. Zani: *Managing Computer-Based Information Systems,* Richard D. Irwin, Inc., Homewood, Ill., 1971.

Holstein, William K.: "Production Planning and Control Integrated," *Harvard Business Review,* May–June 1968, pp. 121–140.

Hoos, Ida R.: "Information Systems and Public Planning," *Management Science,* vol. 17, no. 10, June 1971, p. B-658.

Little, John D. C.: "Models and Managers: The Concept of a Decision Calculus," *Management Science,* vol. 16, no. 8, April 1970, p. B-466.

Mardick, Robert G., and Joel E. Ross: *Information Systems for Modern Management,* Prentice-Hall, Inc., Englewood Cliffs, N.J., 1971.

Meyers, Charles A. (ed.): *The Impact of Computers on Management,* The M.I.T. Press, Cambridge, Mass., 1967.

Ross, Joel E.: *Management by Information System,* Prentice-Hall, Inc., Englewood Cliffs, N.J., 1970.

Whisler, Thomas L.: *The Impact of Computers on Organizations,* Praeger Publishers, New York, 1970.

Zani, William M.: "Real-Time Information Systems: A Comparative Economic Analysis," *Management Science,* vol. 16, no. 6, February 1970, p. B-350.

appendix A

Linear Programming and Simulation

INTRODUCTION

In a course which is applications-oriented it is quite tempting to avoid the detail of the quantitative techniques and to delve right into their application. But before quantitative techniques can be applied, the basic mathematical ideas which underlie them must be studied. Although it is unlikely that any practitioner will go through a solution process manually, there is no substitute for cranking out problems by hand in the initial learning stages to give the confidence and maturity necessary to become a skillful user of a technique. Once this phase has been completed, the computer can be used and the solution of the model can be reduced to a routine procedure.

Two techniques which will be used throughout the text include

1. Linear programming, and
2. Simulation

Although they are not the only ones which we will use, they will serve as a foundation upon which other techniques will be built. It is necessary, then, that the student who has never been exposed to these topics study them carefully. Those who have may benefit from a quick review.

MATHEMATICAL PROGRAMMING

From an overall point of view, the problem faced by the firm is one of constrained maximization. That is, given an objective such as the maximization of profits, the decision maker must decide which products to include in the product line and how many of each to manufacture.

In specifying these levels of output over a planning horizon, the decision maker is constrained by many factors in the environment. In the absence of these constraints, output could be set infinitely large for each product in the line. However, there are usually several constraints operating to prevent this. First, there is a capital constraint—the firm always has a limited budget which can be allocated among several alternatives. Secondly, there are manpower constraints—there is a work force at the beginning of the planning period which constrains output; however, slight changes can normally be made in the work force levels. Third, there are capacity constraints; especially in the short-term and intermediate-term planning periods, the physical facilities are relatively fixed. Finally, one could add social and legal constraints to these economic constraints; however, in many cases they might be difficult to quantify.

The constraints, then, create the environment within which decision makers strive to maximize their objectives. This is the essence of mathematical programming: constrained maximization. It becomes an intuitively appealing framework for the analysis of many types of business problems. The difficult task, however, is shouldered by the model builder, who must abstract from the environment those important elements that are to be incorporated in the model.

A PROBLEM IN CONSTRAINED MAXIMIZATION: GRAPHICAL LINEAR PROGRAMMING

Let us look at a problem faced by a manufacturer of snowmobiles and golf carts. Assuming that they can sell as many as they make, output quantities for both items must be set by management. First, the objective must be set. Usually profit maximization is selected. Second, the constraints that affect the output decisions must be specified. So that the introduction to mathematical programming can be kept at a manageable level, let us assume that there are but two constraints in the manufacturing process. Both constraints will be in the number of man-hours of capacity found in the fabricating and painting departments. Each product must go through both departments. Finally the time horizon over which the plan will be developed must be determined. In this example the horizon will be one month.

Objective The profit contribution of snowmobiles is estimated to be $200 per unit, while that of golfcarts is $300 per unit. Profit contribution is the difference between selling price and variable costs, and it is assumed to remain constant; that is, the first snowmobile sold generates $200 profit and the ten-thousandth

Table A-1 Substitution matrix

	Snowmobiles (S)	*Golfcarts* (G)	*Capacity (in 000's)*
Fabricating (F)	2 hours/unit	1 hour/unit	1,000 man-hours
Painting (P)	1	3	1,200

unit sold generates $200 profit. Because of this property, the objective is called *linear* and can be written as

$$\text{Maximize } P = 200S + 300G$$

where P = profits
S = snowmobiles
G = golfcarts

This equation, which we hope to maximize, is called the *objective function.*

Constraints The capacity next month in the fabricating shop is 1,000 man-hours, while that in painting is 1,200 man-hours. These figures are found in the last column of Table A-1. The remaining entries in the table represent substitution rates. For example, each snowmobile manufactured requires 2 man-hours of fabricating capacity and 1 man-hour of painting capacity. Each golfcart requires 1 man-hour of fabrication and 3 man-hours of painting. The figures in these columns, therefore, represent the inputs necessary for one unit of output.

The limitation on the number of snowmobiles and golfcarts that can be produced in the fabricating department can be expressed mathematically as

$$2S + 1G \leq 1{,}000$$

If no golfcarts were produced, up to 1,000/2 = 500 snowmobiles could be fabricated. On the other hand, if no snowmobiles were produced, up to 1,000/1 = 1,000 golfcarts could be fabricated. It is also possible to fabricate any combination of these two products as long as the constraint is not violated.

The mathematical expression cannot be called an equation, since the "less than or equal to" sign ($\leq$) specifies that the left-hand side must be less than or equal to the right-hand side. Instead, the expression is called an *inequality* and is graphed in Fig. A-1. The intersection with the snowmobiles (S) axis

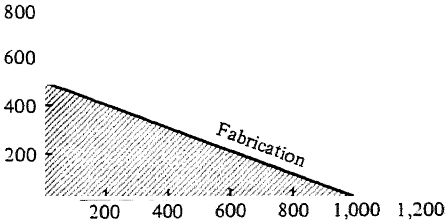


Fig. A-1

represents the 500 snowmobiles that could be made in the fabricating department if no golfcarts were produced. The intersection with the golfcart axis (G) represents the 1,000 golfcarts that could be produced in the fabrication shop if no snowmobiles were produced. Plotting these two extreme points and connecting them with a straight line is the way to start the graphical analysis. Any point on the line represents a combination of golfcarts and snowmobiles which will exactly use up the capacity in the fabrication department. The region which is shown shaded in Fig. A-1 represents those combinations of snowmobiles and golfcarts that will not violate the fabricating capacity.

Next we formulate the constraint for the painting department as

$$1S + 3G \leq 1{,}200$$

It is plotted in Fig. A-2. The shaded area represents combinations of snowmobiles and golfcarts which will not violate the constraints in the painting department.

Feasible region In Fig. A-3 both constraints are plotted. Now it can be seen that when both constraints are taken into consideration, only the cross-hatched area, or that area which is common to both Figs. A-1 and A-2, is feasible. For example, point A, which represents 100 snowmobiles and 600 golfcarts, can be accommodated in the fabrication department but exceeds the capacity of the painting department. Point B, on the other hand, with 200 snowmobiles and 200 golfcarts, can be accommodated in both departments. The shaded region is called the *feasible* region, and the optimum solution will occur at any point within or on the border of this region.

Solution Just where the optimum will occur depends on financial variables. Up to this point only technical variables have been considered. Therefore, the next step is to examine the role of profit contribution in determining the best point in the feasible region.

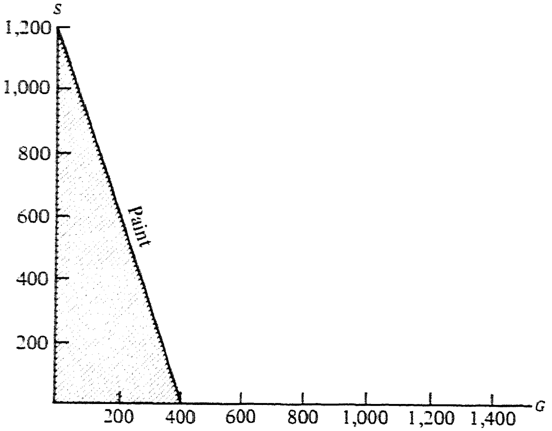


Fig. A-2

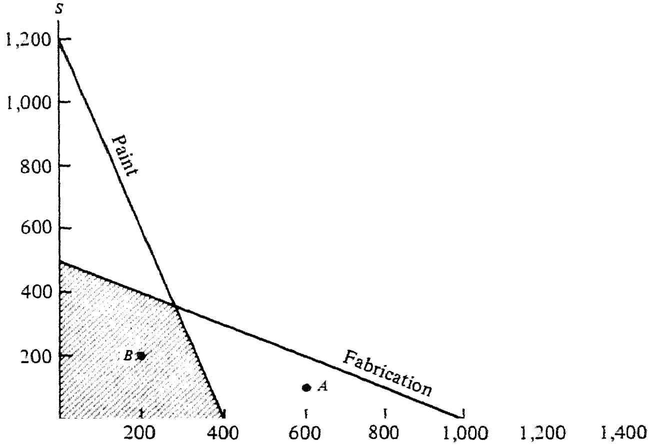


Fig. A-3

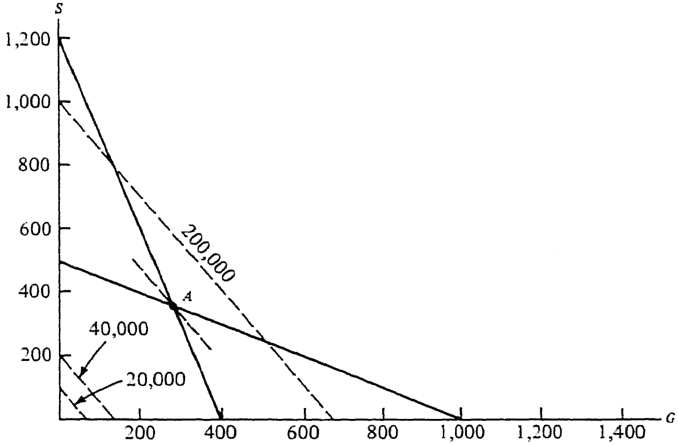


Fig. A-4

To accomplish this, we make use of the objective function

$$P = 200S + 300G$$

to establish a series of iso-profit lines. For example, take any arbitrary output of snowmobiles and no golfcarts. If we produce 100 snowmobiles, profit is $200(100) = \$20{,}000$. Marking this point on the S axis in Fig. A-4 establishes one point on a \$20,000 iso-profit line. The other point on the G axis can be determined by finding how many golfcarts must be sold to generate \$20,000 of profit. The answer is $20{,}000/300 = 66$ golfcarts. The line connecting these two points represents all combinations of snowmobiles and golfcarts that will generate \$20,000 worth of profit. Profit lines can also be determined for \$40,000, \$80,000 \$100,000, and so on. In fact all iso-profit lines will be parallel to the first one drawn. Furthermore, iso-profit lines generating higher profits always lie to the right of lower ones.

Does the \$20,000 iso-profit line represent a feasible solution? Yes. All points on the \$20,000 line lie inside the feasible region. Why not operate somewhere on this line? Because we have the capacity to operate on a higher iso-profit line to the right. Problems occur, however, if we move as far to the right as the \$200,000 profit line. Nowhere on this line do we have enough capacity to turn out combinations of snowmobiles and golfcarts which would generate \$200,000 profit. The solution, then, is to operate on that iso-profit line which is furthest out yet does not violate any of the constraints. *A* in Fig. A-4 is just this point. Reading from the graph, the number of snowmobiles that should be produced is about 360 and the number of golfcarts is about 280. Notice that the solution occurs at a corner; in fact, the solution will *always* occur at a corner. There is the special case, however, where the iso-profit line is parallel to one of the constraints. In this situation there will be an infinite number of solutions along this line, including, of course, the two corners.

If there had been more departments, the only changes would have been the addition of constraints. A smaller feasible region would have resulted. If, however, there are more than two products, we have troubles. Suppose that the manufacturer produces snowmobiles, golfcarts, outboard engines, and lawn mowers. Solving the problem would require us to work in four-dimensional space. Since this is an impossibility, another approach is required for these more complex situations.

SIMPLEX METHOD

The simplex method, unlike the graphical, can handle any number of products or activities. So that an intuitive understanding of the simplex method can be grasped, the problem just solved graphically will be solved using the simplex method. The problem is written in its inequality form as

$$\begin{aligned} &\text{Maximize } P = 200S + 300G \\ &\text{S.T.} \qquad\qquad 2S + 1G \leq 1{,}000 \\ &\qquad\qquad\qquad\; 1S + 3G \leq 1{,}200 \end{aligned}$$

The "S.T." specifies that the objective function is "subject to" the constraints which follow.

Augmented form In order to avoid working with inequalities, it is necessary to make some changes in the constraint set. The changes, however, must still preserve the relationship between the left- and right-hand side of the inequality sign. This can be accomplished by introducing a slack variable. For example, the first inequality in the constraint set can be written as

$$2S + 1G + L_1 = 1{,}000$$

where L_1 = slack in fabricating department

Slack is zero if the number of snowmobiles plus the number of golfcarts produced in the fabrication department uses up all the capacity. If they do not, L_1 will take on whatever value is necessary to absorb the excess. For example, if 200 snowmobiles and 200 golfcarts are produced, $L_1 = 400$. The second constraint equation can now be written as

$$1S + 3G + L_2 = 1{,}200$$

Since two new variables have been introduced into the constraint set, it is necessary that they appear in the objective function. What will the coefficients of these slacks be in the objective function? A slack contributes nothing to profit; therefore the objective function can be written as

$$\text{Maximize } P = 200S + 300G + 0L_1 + 0L_2$$

The new set of equations, including objective function and constraints, are now said to be in "augmented form." Rewriting them, we have

$$\begin{array}{llcl} \text{Maximize } P = & 200S + 300G + 0L_1 + 0L_2 & & \\ \text{S.T.} & 2S + \quad 1G + 1L_1 & & = 1{,}000 \\ & 1S + \quad 3G & + L_2 & = 1{,}200 \end{array}$$

Initial matrix For ease in manipulating the equations, a matrix is set up. Across the top row are placed all the variables as we have them in the objective function of the augmented form. See Table A-2. Down the second column we place the initial set of variables that is produced. A characteristic of the simplex technique is that in this first step *only* the *slack* variables are produced. That is, we start at a point where the output of the non-slacks is zero. Into the last column goes the quantity of each slack produced, and in this first step the quantity is equal to the total supply of each resource. Next we enter the substitution rates in their appropriate locations within the matrix. Finally the profit contributions are entered above the variables and alongside the slacks in the first column.

Intuitive explanation of the simplex method Before presenting the mathematical solution process, a verbal description of this process will be given so that an intuitive feel can be obtained for the problem.

Table A-2

		200	*300*	*0*	*0*	
		S	G	L_1	L_2	
0	L_1	2	1	1	0	1,000
0	L_2	1	3	0	1	1,200

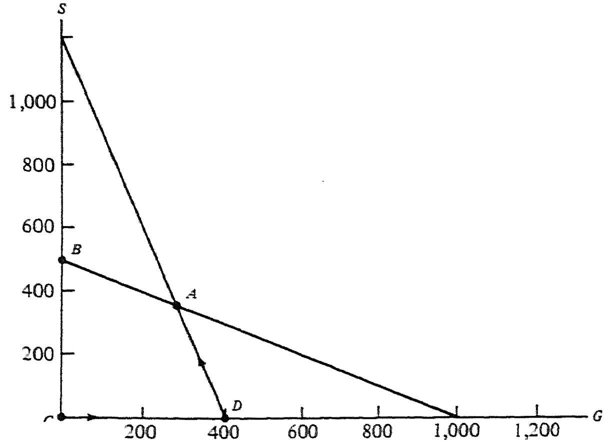


Fig. A-5

Looking at Fig. A-5, you can see that the solution process starts at the origin, where neither snowmobiles nor golfcarts are produced. At this point slack in the fabrication department L_1 is 1,000 and slack in the painting department L_2 is 1,200. The procedure then determines the most profitable direction in which to head. If we start producing just snowmobiles, for every unit sold a profit of $200 is generated; whereas if golfcarts are produced, a profit per unit of $300 is added. Therefore the optimal strategy would be to add as many golfcarts as possible. Golfcarts are then added until a constraint is violated. In this case we can see from Fig. A-5 that painting capacity runs out at point D.

At point D the same kind of question will be asked: Would it be profitable to proceed in the direction of point A? Since we already know from our graphical analysis that point A is the optimal solution, it seems reasonable for the simplex method to move us from D toward A. How far do we move? Again, until a constraint is hit—point A. We again ask if it will be profitable to move on the new constraint toward B. The answer we would expect is, no. That is, A is the best point.

First iteration Getting back to Table A-2, we can see how this procedure is executed mathematically. The first question that must be answered is, Which variable comes into solution? In this case is it S or G? Each must be evaluated separately, and the net effect of the change must be considered. For example, the net effect of bringing in S on a per-unit basis can be evaluated from the S column in Table A-2.

$$\begin{array}{r} +200 \\ -2(0) \\ \underline{-1(0)} \\ +200 \end{array}$$

First, the addition of 1 unit of S will generate $200. Next, 2 units of slack in the fabricating department must be given up; and since it contributes nothing to profit, we write its impact on profit contribution as —2(0). Mechanically the 2 is found at the intersection of the L_1 row and the S column. The zero is found at the outside left-hand corner of the L_2 row. Table A-2 also tells us that 1 unit of slack in the painting department will be given up; and since slack here contributes nothing to profit, we have —1(0). The net effect is the sum of these, or $200.

Now the per-unit consequence of introducing G must be evaluated. It can be computed as

$$\begin{array}{l} +300 \\ -1(0) \\ \underline{-3(0)} \\ +300 \end{array}$$

The introduction of G adds $300 to profits. However, one unit of slack in the fabricating department, which contributes nothing to profits, must be given up. In addition, 3 units of slack in the painting department, also contributing nothing to profits, must be given up. The net contribution is $300 per unit. It is not necessary to compute the consequence of introducing L_1 and L_2, since they are already in the solution.

The decision as to which variable enters the solution can now be made. It is that variable with the largest per-unit profit contribution. In this case it is G, with a per-unit net contribution of $300, just as we predicted in the intuitive development. The next problem is to determine just how much of G should be brought in.

Every unit of G brought into the solution requires that 1 unit of slack in fabricating and 3 units of slack in painting be consumed. With only 1,000 units of capacity in fabricating, the maximum number of G's that can be brought into the fabricating department is

$$\frac{1{,}000}{1} = 1{,}000$$

With 1,200 units of capacity in the painting department, the maximum number of G's that can be brought in is

$$1{,}200 = 400$$

We now have two limits to consider. In the fabricating department the limit is 1,000 units of G and in the painting department the limit is 400 units of G. The smaller of the two limits our choice, since it would be senseless to send 1,000 units through fabrication if we could paint only 400 of them. The choices made up to this point are indicated by arrows in Table A-3.

Relating the progress made to Fig. A-5, we started with a matrix describing the situation at point C. We then proceeded to move in the G direction and

Table A-3

		200	*300*	*0*	*0*			
		S	G	L_1	L_2			
0	L_1	2	1	1	0	1,000	$\frac{1,000}{1}$	1,000
0	L_2	1	3	0	1	1,200	$\frac{1,200}{3}$	400

Table A-4

		200	*300*	*0*	*0*	
		S	G	L_1	L_2	
0	L_1	$\frac{5}{3}$	0	1	$-\frac{1}{3}$	600
300	G	$\frac{1}{3}$	1	0	$\frac{1}{3}$	400

discovered that the first constraint met was at point D. The next step is to develop a new matrix describing the situation at point D.

First, a new matrix outline is drawn. See Table A-4. Along the top row the variables are entered just as they were in the first matrix with their profit contributions above them. The produced variables, however, are L_1 and G. Since L_2 was reduced to zero (we used up all the slack in painting), it no longer appears as a produced variable. In its place we have G. In fact, when one variable is entered, one leaves. Therefore, if there are two rows of numbers in the first matrix, there will always be two rows in the next and subsequent matrices. Since we did not use up all of the capacity in the fabrication shop, L_1 still appears as a produced variable.

What are the produced amounts of L_1 and G? The value of G is 400 units, as we just determined in Table A-3. The value of L_1, however, has not been computed as yet. From Table A-3 it can be seen that we started with 1,000 units of L_1. For every unit of G brought into the solution, our substitution matrix in Table A-3 tells us that 1 unit of L_1 was required. Since 400 units of G were brought in, we have

$$1{,}000 - 400(1) = 600$$

units of L_1 remaining. The 600 is then entered in the last column of Table A-4. It now remains to calculate the new substitution rates.

Starting with the variable which has just been introduced, G, we write its column equation from Table A-3 as

$$G = 1L_1 + 3L_2$$

This is an equation with L_1 and L_2 on the right-hand side, and our new matrix requires L_1 and G on this side. That is, the new produced variables must appear on the right-hand side. To accomplish this we solve this equation for L_2.

$$3L_2 = G - L_1$$

$$L_2 = \tfrac{1}{3}G - \tfrac{1}{3}L_1$$

We can now enter our new column equation for L_2 in Table A-4.

Next the column equation for S is written from Table A-3:

$$S = 2L_1 + 1L_2$$

Again L_2 is on the right; and before the equation can be entered in Table A-4, it must have only G and L_1 on the right-hand side. Since we have already determined a new equation for L_2 in the previous step, we can substitute this into the equation for S.

$$S = 2L_1 + 1(\tfrac{1}{3}G - \tfrac{1}{3}L_1)$$
$$S = 2L_1 + \tfrac{1}{3}G - \tfrac{1}{3}L_1$$
$$S = \tfrac{5}{3}L_1 + \tfrac{1}{3}G$$

Now we have an equation for S in only L_1 and G, and it can be entered in Table A-4. The procedure of substituting the equation

$$L_2 = \tfrac{1}{3}G - \tfrac{1}{3}L_1$$

for the variable that has just left solution, L_2, is continued until all the remaining column equations have G and L_1 on the right-hand side.

This process, however, stops here in this example, for the only remaining columns that have not been completed are G and L_1, and these are the produced variables in our new table. Wherever the column variables are the same as the produced variables, a "1" is entered in the matrix where they intersect. Zeros comprise the remaining entries in those columns. Therefore, we enter a "1" where the column and row G intersect, a zero above it, a "1" where the column and row L_1 intersect, and a zero below it. Prove that these column entries are correct by substituting the new equation for L_2 in these column equations of Table A-3.

Now we are at point D in Fig. A-5, and the output levels can be read from Table A-4.

$$L_1 = 600$$
$$G = 400$$

One iteration has now been completed, and the remainder of the analysis is executed exactly like the first. To continue, we must ask if it is profitable to proceed toward point A in Fig. A-5. If so, how far? This is the objective of the second iteration.

Second iteration As in the first iteration, the first step is to evaluate the consequence of bringing a new variable into the solution. Working from Table A-4, we have

S	L_2
$+200$	$+0$
$-\frac{5}{3}(0)$	$-(-\frac{1}{3})(0)$
$-\frac{1}{3}(300)$	$-\frac{1}{3}(300)$
$+100$	-100

The net contribution of bringing in S is made up of adding one unit of S, which contributes \$200 to profit, and taking out $\frac{1}{3}$ of a unit of G, which reduces contribution by $\frac{1}{3}(300) = \$100$, for a net contribution of \$100. It can be seen from Fig. A-5 that if one goes from point D to A, the introduction of some snowmobiles must be accompanied by the reduction in the number of golfcarts. The net contribution associated with bringing in L_2 is —\$100; and since the activity with the highest net contribution is introduced, we choose S.

Exactly how much of S is introduced is determined as before. From Table A-5 it can be seen that 600 units of L_1 remain; and for every unit of S introduced, $\frac{5}{3}$ of a unit of L_1 must be given up. Then $600/(\frac{5}{3}) = 360$ of S can be introduced. The same analysis holds for G. Since 400 units of G are produced, and for every S introduced $\frac{1}{3}G$ must be given up, $400/\frac{1}{3} = 1{,}200$ units of S can be introduced. Again we choose the smallest of the two, or 360, since it makes no sense to introduce 1,200 units of S if only 360 can be processed through the fabrication shop. Since 360 units of S will be introduced, all the capacity in L_1 will be exhausted, and we will therefore have as new produced variables S and G.

A new matrix is drawn in Table A-6 and the new produced variables S and G can be entered in the second column. The produced amount of the variable S is 360, as was just determined. The produced amount of G can be

Table A-5

		200	*300*	*0*	*0*		
		S	G	L_1	L_2		
0	L_1	$\frac{5}{3}$	0	1	$-\frac{1}{3}$	600	$600/\frac{5}{3} = 360$
300	G	$\frac{1}{3}$	1	0	$\frac{1}{3}$	400	$400/\frac{1}{3} = 1{,}200$

Table A-6

		200	*300*	*0*	*0*	
		S	G	L_1	L_2	
200	S	1	0	$\frac{3}{5}$	$-\frac{1}{5}$	360
300	G	0	1	$-\frac{1}{5}$	$\frac{6}{15}$	280

calculated from Table A-5, where the level of G was 400, and it is shown that for every S introduced $\frac{1}{3}$ of a unit of G is given up. Since 360 units of S will be introduced, the level of G in the new matrix will be

$$G = 400 - \tfrac{1}{3}(360) = 280$$

This is entered in the last column of the matrix in Table A-6. The substitution rates for the variable introduced, S, can now be computed as:

$$S = \tfrac{5}{3}L_1 + \tfrac{1}{3}G$$

Since we now require an S and G on the right-hand side, we solve for L_1:

$$\tfrac{5}{3}L_1 = S - \tfrac{1}{3}G$$

$$L_1 = \tfrac{3}{5}S - \tfrac{1}{5}G$$

The column equation for L_2 can be written as

$$L_2 = -\tfrac{1}{3}L_1 + \tfrac{1}{3}G$$

It also must have S and G on the right-hand side. Therefore we substitute the new value for L_1 and get

$$L_2 = -\tfrac{1}{3}(\tfrac{3}{5}S - \tfrac{1}{5}G) + \tfrac{1}{3}G$$

$$L_2 = -\tfrac{1}{5}S + \tfrac{1}{15}G + \tfrac{1}{3}G$$

$$L_2 = -\tfrac{1}{5}S + \tfrac{6}{15}G$$

The new matrix represents conditions at point A. The output levels can be read from the matrix as

$$S = 360$$

$$G = 280$$

Third iteration The process repeats itself until it is no longer profitable to introduce new variables into the solution. First the consequence of bringing new variables into the solution must be considered.

L_1	L_2
$+0$	$+0$
$-\frac{3}{5}(200)$	$-(-\frac{1}{5})(200)$
$-(-\frac{1}{5})(300)$	$-\frac{6}{15}(300)$
-60	-80

Since the net contributions come out negative, the introduction of either L_1 or L_2 will reduce profits. Therefore we go no further and use the results of the second iteration as the optimal intensity of operation. To maximize profit, we produce:

$$S = 360$$

$$G = 280$$

The profit generated will be

$$200(360) + 300(280) = \$156{,}000$$

Summary of simplex

Step 1: Which variable enters?
Step 2: The maximum that can be brought into the solution.
Step 3: Draw new matrix outline.
Step 4: Enter produced variables.
Step 5: Enter amounts produced.
Step 6: Determine new substitution rates.
Step 7: Repeat process, starting at step 1.

Shadow prices The net contributions in the solution basis, —60 and —80, have a special meaning. They represent the increase in profit that would result if an additional unit of scarce resource were obtained. In this example if an additional unit of fabrication were obtained, profit would increase by \$60. If an additional unit of painting capacity were obtained, profit would increase by \$80. This tells the decision maker the maximum price he should be willing to pay if he had the opportunity to buy additional units of either scarce resource. Therefore, if someone offered him additional fabricating capacity at \$50 per unit, he should accept. At \$60 he is indifferent, and above this point his costs are greater than his benefits. His point of indifference for painting capacity is \$80.

These special net contributions associated with slack variables are called *shadow prices* and impute a value to scarce resources.

Now that you have gone through the process once, write out the example which follows and try to solve it. Immediately following the statement of the problem is its solution. Use it only if you get stuck or to check your answer.

Example:

$$\begin{aligned} \text{Maximize } P &= 8v + 17x + 14y \\ \text{S.T.} \quad & 3v + 7x + 4y \le 70 \\ & 2x + 4y \le 80 \\ & v, x, y \ge 0 \end{aligned}$$

Solution: Written in the augmented form, we have

$$\begin{aligned} \text{Maximize } P &= 8v + 17x + 14y + 0L_1 + 0L_2 \\ \text{S.T.} \quad & 3v + 7x + 4y + 1L_1 = 70 \\ & 2x + 4y + 1L_2 = 80 \end{aligned}$$

		8	17	14	0	0	
		v	x	y	L_1	L_2	
0	L_1	3	7	4	1	0	70
0	L_2	0	2	4	0	1	80

Which variable enters?

v	x	y
8	17	14
−3(0)	−7(0)	−4(0)
−0(0)	−2(0)	−4(0)
8	17	14

Maximum that can be brought into the solution:

$$\frac{70}{7} = 10$$

80

$= 40$

Amounts produced:

$x = 10$

$L_2 = 80 - 2(10) = 60$

New substitution rates:

$x = 7L_1 + 2L_2$

Solving for L_1:

$$\boxed{L_1 = \tfrac{1}{7}x - \tfrac{2}{7}L_2}$$

$v = 3L_1$

$v = 3(\tfrac{1}{7}x - \tfrac{2}{7}L_2)$

$v = \tfrac{3}{7}x - \tfrac{6}{7}L_2$

$y = 4L_1 + 4L_2$

$y = 4(\tfrac{1}{7}x - \tfrac{2}{7}L_2) + 4L_2$

$y = \tfrac{4}{7}x + \tfrac{20}{7}L_2$

		8	*17*	*14*	*0*	*0*	
		v	x	y	L_1	L_2	
17	x	$\frac{3}{7}$	1	$\frac{4}{7}$	$\frac{1}{7}$	0	10
0	L_2	$-\frac{6}{7}$	0	$\frac{20}{7}$	$-\frac{2}{7}$	1	60

Which variable enters?

L_1

$$\begin{array}{r} 8 \\ -\dfrac{3}{7}\,(17) \\ -\left(-\dfrac{6}{7}\right)(0) \\ \hline +\dfrac{5}{7} \end{array} \qquad \begin{array}{r} 14 \\ -\dfrac{4}{7}\,(17) \\ -\dfrac{20}{7}\,(0) \\ \hline +\dfrac{30}{7} \end{array} \qquad \begin{array}{r} \\ -\dfrac{1}{7}\,(17) \\ -\left(-\dfrac{2}{7}\right)(0) \\ \dfrac{17}{7} \end{array}$$

Maximum that can be brought into the solution:

$$\frac{10}{\frac{4}{7}} = \frac{70}{4} = \frac{350}{20} \quad \longleftarrow$$

$$\frac{60}{\frac{20}{7}} \qquad \frac{420}{20}$$

Amounts produced:

$$y = \frac{70}{4}$$

$$L_2 = 60 - \frac{20}{7}\left(\frac{70}{4}\right) = 10$$

New substitution rates:

$$y = \tfrac{4}{7}x + \tfrac{20}{7}L_2$$

$$x = \tfrac{7}{4}y - 5L_2$$

$$v = \tfrac{3}{7}x - \tfrac{6}{7}L_2$$
$$v = \tfrac{3}{7}(\tfrac{7}{4}y - 5L_2) - \tfrac{6}{7}L_2$$
$$v = \tfrac{3}{4}y - 3L_2$$
$$L_1 = \tfrac{1}{7}x - \tfrac{2}{7}L_2$$
$$L_1 = \frac{1}{7}\left(\frac{7}{4}\,y - 5L_2\right) - \frac{2}{7}\,L_2$$

$$L_1 = \frac{}{4}\,y - L_2$$

		8	17	14	0	0	
		v	x	y	L_1	L_2	
14	y	$\frac{3}{4}$	$\frac{7}{4}$	1	$\frac{1}{4}$	0	$\frac{70}{4}$
0	L_2	-3	-5	0	-1	1	10

Which variable enters?

v	x	L_1
8 $-\frac{3}{4}(14)$ $-(-3)(0)$	17 $-\frac{7}{4}(14)$ $-(-5)(0)$	0 $-\frac{1}{4}(14)$ $-(-1)(0)$
$-\frac{10}{4}$	$-\frac{30}{4}$	$-\frac{14}{4}$

Optimal solution is:

$$y = \frac{70}{4} \qquad v = 0 \qquad L_1 = 0$$

$$L_2 = 10 \qquad x = 0$$

$$\text{Profit} = 8(0) + 17(0) + 14\left(\frac{70}{4}\right) = 245$$

Sensitivity analysis In many cases the most useful aspect of a programming model is more than just the optimum solution. In addition it is the way in which changes in the parameters can affect this solution. For example, it is useful to know how changes in profit contributions, capacities, constraint coefficients, and so on, will affect the optimum solution. One approach would be to make the changes in the appropriate parameters and solve the problem all over again. If the optimum solution did not change or did not change by very much, it can be concluded that within the range investigated the solution is insensitive to changes. This type of analysis, which tests the sensitivity of the optimum solution to changes in parameters, is called *sensitivity analysis*.

BASIC ASSUMPTIONS OF LINEAR PROGRAMMING

Linear programming analysis assumes that the relationships in the model, both objective function and constraint set, are linear. In addition, the coefficients in all these equations are assumed to be known with certainty. There seems to be something artificial when the equations are written with just one number as a coefficient. Does it always take exactly 2 man-hours per golfcart in the fabrication shop? Chances are that sometimes it takes more and sometimes less. The question, then, is whether we miss something in our analysis by avoiding this variability.

Linearity and certainty, then, are two major assumptions which must be kept in mind when the linear model is considered. One should always ask, Is the relationship sufficiently linear within the range of interest? Is it reasonable to assume certainty? In situations where these are not reasonable assumptions to make, other programming formulations must be considered. These other techniques can be classified as

1. Nonlinear programming
2. Integer programming
3. Stochastic programming

Although a full development for each of these techniques is beyond the scope of this book, a few paragraphs on each might prove helpful in understanding when these advanced methods would be employed. Before we take off on this tangent, it is important to recognize that linear programming is still the most popular. In many cases a linear approximation to a problem proves to be adequate.

INTEGER, NONLINEAR, AND STOCHASTIC PROGRAMMING

Integer programming In the linear programming example, the solution came out as integer values. However, we were lucky. There was nothing in the specification of the problem that insisted on integer values. It is quite likely, given other values as coefficients, that the outcome could have been to produce 274.5 snowmobiles and 363.7 golfcarts. Simply rounding these values off to the nearest integer might violate the constraints. In fact, going from the linear solution to an integer solution is usually not a trivial task. Many of the techniques that are used are beyond the scope of this book; however, in some cases where integer solutions are absolutely necessary a simple technique for finding integer solutions will be presented. In most other cases it will be sufficient to use linear approximations and round off.

Nonlinear programming In the case where either the objective function or the constraints are nonlinear, use is made of nonlinear programming techniques. Nonlinearities exist, for example, when a firm must lower prices to induce more people to buy, or when the firm is in a region of increasing or decreasing marginal costs.

In many of these nonlinear situations, linear programs have been used as approximations and have proved to be adequate. However, with the increase in sophistication of the users, together with availability of nonlinear computer routines, there will be a marked increase in the use of this technique.

Using intuition, it is possible to see the consequence of nonlinear programming techniques. First let us explore a special characteristic of a linear model and determine what we would expect to happen in the nonlinear case. In the linear formulation we saw from our example that the number of nonzero activities, or products, in the final solution was equal to the number of constraints: two. In fact, this relationship is always true; the number of activities in the final solution will equal the number of constraints. Surprisingly, then, if a firm produces 2,000 products and the products must go through two manufacturing departments, the firm can maximize profits by producing but two products.

If the profit were subject to diminishing marginal profitability, where the price must be lowered to induce higher sales, it would be reasonable to expect that there will be more than two products included in the solution set. Fewer of each product would be sold. Therefore in nonlinear programming with an objective function displaying diminishing returns, we would expect more nonzero activities in the final solution than the number of constraints. If, on the other hand, we were in a region of increasing returns, we would expect fewer nonzero activities in the solution set. Explain why.

Stochastic programming In the linear programming example, we assumed that profit contributions, substitution rates, and capacities were all known with certainty. It is very rare, however, that anything is known with certainty. Suppose that production costs varied within some range, thereby making the profit contribution of each product a variable. It is also likely that capacity may not be known with certainty. In fact, variability is a characteristic of nature. In many cases these uncertainties significantly influence the way in which the model is built, the way in which it is solved, and the solution itself. When we wish to take explicit account of all the possible values of the variables, many types of stochastic models are available.

The programming exercises in this book, however, will use average values rather than the full range of possible outcomes. For example, it is implied that the average profitability of golfcarts is $300. In many cases this certainty approximation to the stochastic case is quite adequate.

SIMULATION

Simulation is the duplication of the essence of a system without actually obtaining reality. Usually the simulation is a more complex and more realistic representation of reality than is a linear programming model. In addition, a simulation model, unlike linear programming, takes uncertainty into account. There is a danger, however, in building models that are so complex that they begin to resemble reality itself. The danger is that the very insight which the model was supposed to provide might be lost. Therefore the model builder must realize that there is a trade-off when making the model complex. He must ask himself whether he needs to deal with the full complexity of the situation to gain useful insight.

OBTAINING INPUT DATA

Consider a situation where customers queue up at a bank window for service. Two main factors influence the way in which a queue builds up. The first is the time pattern of arrivals at the window, while the second is the time for servicing an arrival. By observing these arrival and service rates, it will be possible to simulate the queue.

To determine the behavior of arrivals, a sample of 100 customers is taken and the time between arrivals is recorded in Table A-7. It can be seen that the time between arrivals for 10 of the observations was 1 minute, for 15 of the observations 2 minutes, and so on.

Table A-7 Time between arrivals

Time between arrivals (min)	*Number of occurrences*
1	10
2	15
3	40
4	20
5	15

Table A-8 Service time

Service time	*Number of observations*
1	10
2	25
3	40
4	15
5	10

Observing the service times for 100 customers, we record the results in Table A-8. On 10 occasions a service time of 1 minute was recorded, 25 times a service time of 2 minutes was recorded, and so on.

PROCESS OF SIMULATION

Now an arrival pattern can be simulated at the teller's cage such that it maintains the same statistical characteristics as the observed arrival pattern. To accomplish this we will generate arrival times so that they have the same probability of occurring in our simulation experiment as they did in our sample. That is, a time between arrival of 1 minute will have a 10 percent chance of occurring, a time between arrival of 2 minutes will have a 15 percent chance, and so on. Using a brute force approach, this could be accomplished by tearing up 100 pieces of paper and on 10 of them writing "1 min," on 15 of them writing "2 min," on 40 of them writing "3 min," and so on. Next, one would place the bits of paper in a hat and randomly pick out pieces, thereby generating an arrival pattern. For example, the first piece of paper selected might be "3 min." If the bank opens at 9:00 A.M., this implies that the first arrival is at 9:03. If the second draw from the hat yields "2 min," this indicates that the next arrival occurs 2 minutes *later,* or at 9:05. This process could continue until the arrivals for an entire day are simulated.

Use of random numbers Since picking numbers at random out of a hat is a bit cumbersome (especially for a computer), random numbers are usually used to sample from a distribution such as ours. To accomplish this, first a series of two-digit numbers are assigned to each set of arrival times. The two-digit numbers are in sequential order, and the quantity of them assigned corresponds to the number of occurrences observed. The numbers are assigned in Table A-9.

Now a series of random numbers is generated from a random number table (see Appendix B). Selection of the numbers can proceed from any place in the table. For example, one might start down the first column, reading the first two digits of each number, since a two-digit random number is required for this problem. A series of numbers is recorded in the first column of Table A-10. Using these numbers, the corresponding times between arrivals can be

Table A-9

Time between arrivals	*Number of occurrences*	*Two-digit numbers*
1	10	00–09
2	15	10–24
3	40	25–64
4	20	65–84
5	15	85–99

Table A-10

Random number	*Corresponding time between arrivals*	*Clock time at arrival*
10	2	9:02
37	3	9:05
08	1	9:06
99	5	9:11
12	2	9:13
66	4	9:17
31	3	9:20
85	5	9:25
63	3	9:28
73	4	9:32

Table A-11

Service time	*Number of observations*	*Corresponding two-digit number*
1	10	00–09
2	25	10–34
3	40	35–74
4	15	75–89
5	10	90–99

read from Table A-9. Both the times between arrivals and the clock times are shown in Table A-10. To keep the calculations manageable, only a half-hour is simulated.

The next step is to generate a series of service times. As was done for the arrival rates, a set of corresponding two-digit numbers is assigned to the observed service times. See Table A-11.

Again a series of random numbers is generated and the corresponding service times are identified. The results are shown in Table A-12.

Table A-12

Random number	*Corresponding service time*
98	5
11	2
83	4
88	4
99	5
65	3
80	4
74	3
69	3
09	1

Table A-13

Clock time at arrival	*Can start service*	*Service time*	*Service complete*	*Waiting time*
9:02	9:02	5	9:07	0
9:05	9:07	2	9:09	2
9:06	9:09	4	9:13	3
9:11	9:13	4	9:17	2
9:13	9:17	5	9:22	4
9:17	9:22	3	9:25	5
9:20	9:25	4	9:29	5
9:25	9:29	3	9:32	4
9:28	9:32	3	9:35	4
9:32	9:35	1	9:36	3

Now the simulated arrival rates and service times can be combined to simulate the behavior of the queue for a half-hour period. This final process can be seen in Table A-13. The columns headed "Clock time at arrival" and "Service time" can be completed from Tables A-10 and A-12. Starting at the first arrival, service can begin immediately, since there is no queue. It takes 5 minutes to service this customer; therefore service is complete at 9:07.

The last entry for this customer is the time spent waiting; in this case, it is zero. The next arrival occurs at 9:05. Service, however, cannot commence until 9:07, since service on the first arrival does not terminate until that time. With 2 minutes of service time, service is complete at 9:09, and 2 minutes were spent in line. Applying this same logic, the remainder of the table is filled out. If waiting time is accumulated up to 9:30, 27 minutes were spent in line by customers. If the bank is open 6 hours, one could estimate that 12 × 27 or 324 minutes would be spent waiting in line. Rather than multiplying the results of the half-hour simulation by 12, perhaps a more realistic approach would be to continue the simulation for the 6-hour day.

INDEPENDENCE

One of the basic assumptions in our model is the independence between arrival rates and service times. That is, we have no reason to believe that a long time in between arrivals is associated with a short service time or that a short time between arrivals is associated with a long service time. In fact, we assume that time between arrivals is totally independent of the time for service. Although this condition is a perfectly reasonable assumption in this simulation experiment, it may not always be the case.

THE SAMPLE SIZE FOR EMPIRICALLY DERIVED DISTRIBUTIONS

As data for the simulation, time between arrivals and service times was observed for 100 events. It was these statistical patterns that were used to simulate the process. As is the case whenever sampling takes place, the question that comes up is, Should a larger sample be taken? In this case, should more than 100 arrivals and services have been observed? Perhaps the sample of 100 observations does not give a representative pattern of arrivals and services. The question of sample size is a troublesome one. However, other than bringing up the problem, we refer the reader to more advanced treatments of simulation.

Another problem with our sampling procedure is that it is very likely that the nature of the arrival pattern may change over time—perhaps every hour, or even every 15 minutes. Certainly it might be reasonable to expect that the time between arrivals is less at 11:30 than it is at 9:30. If this is the case, several distributions should be used, one for each period in the day.

HOW MANY REPLICATIONS?

Once the problem of specifying input distributions has been solved, the demands made on the output must be considered. For example, will a simulation of one 6-hour day provide the accuracy necessary for decision making? This is unlikely, since the random generation of arrival times will give an answer in one simulation run that will not be duplicated in the next. Therefore, several replications of the simulation for a 6-hour day must be considered. Exactly how many should there be? Well, if you recall your basic statistics, this problem is faced in any sampling experiment. The larger the sample, or in this case the more the runs, the greater the confidence we have in the results. Usually several simulation runs will be executed, and the mean of these runs is used as the figure of merit. The exact number of runs depends on the accuracy or confidence required.

SIMULATION AS AN OPTIMIZING TECHNIQUE

It is unlikely that the alternative of a single teller represents the optimal solution to the problem. Perhaps an additional teller might make an improvement in operating performance. To examine the consequence of an additional teller, another simulation could be run with the same arrival and service patterns. If a customer arrives when teller A is busy, he can proceed to teller B. Only if both are servicing a customer does a queue start to build. Incorporating this modification, a simulation with two tellers can be carried out. Of course, the total waiting time will be lower, but this has been accomplished at the expense of an additional teller. If management can assess the cost per minute of waiting time, the additional cost of a teller can be compared with the savings in waiting time.

If it pays to install another teller's cage, does this necessarily imply that the optimal solution has been discovered? No. All we have succeeded in doing is comparing two alternatives, the first being one teller and the second two tellers. There are many other simulations that could be run: three, four, five tellers, and so on. In fact, one could even conceptualize new approaches, such as separate tellers for commercial and personal transactions, or separate windows for short transactions perhaps involving less than three items. Each alternative could be simulated and a figure of merit such as net savings could be calculated.

Simulation, then, determines the merit of a particular alternative, rather than determining the best alternative. To be reasonably sure that a near-optimal solution to a particular problem is achieved, many alternatives must be evaluated.

QUESTIONS

QA-1. What is the difference in interpretation between an equality and an inequality? How can an inequality be converted to an equality?

QA-2. Can the firm's objective be considered as one-dimensional (profit)? Is it multidimensional? What about educational institutions, public utilities, and hospitals? If their objectives are multidimensional, what factors beside profit should be considered?

QA-3. Is pollution control a constraint which must be considered in management's quest for constrained optimization? How can it be quantified?

QA-4. What is an iterative procedure?

QA-5. What is the economic interpretation of the shadow price?

QA-6. Define: (*a*) stochastic programming, (*b*) sensitivity analysis, (*c*) integer programming, and (*d*) replication.

QA-7. Describe the basic steps which are necessary in developing and executing a simulation model.

QA-8. How would you design a simulation experiment for determining the adequacy of a new double-chair lift at Mt. Stowe, Vermont? What problems must be resolved? What assumptions would have to be made?

QA-9. Do you think it would be possible to build a model for simulating a football game between the New York Jets and the New England Patriots? If so, how would you go about it? How long would it take you to complete the project?

PROBLEMS

PA-1. Solve the following problem (*a*) graphically and (*b*) by the simplex method.

$$\begin{aligned} \text{Maximize } p &= 10x_1 + 8x_2 \\ &3x_1 + 1x_2 \leq 20 \\ &2x_1 + 3x_2 \leq 10 \\ &x_1,\ x_2 \geq 0 \end{aligned}$$

PA-2. Solve by the simplex method.

$$\begin{aligned} \text{Maximize } p &= 5x_1 + 3x_2 + 1x_3 \\ &1x_1 + 2x_2 + 1x_3 \le 10 \\ &3x_1 + 1x_2 + 2x_3 \le 12 \\ &1x_1 + 1x_2 + 1x_3 \le 7 \\ &x_1, x_2, x_3 \ge 0 \end{aligned}$$

PA-3. Solve the following problem graphically.

$$\begin{aligned} \text{Minimize } c &= 3x_1 + 2x_2 \\ &4x_1 + 6x_2 \ge 3 \\ &7x_1 + 5x_2 \ge 4 \\ &x_1, x_2 \ge 0 \end{aligned}$$

PA-4. Solve by the simplex method.

$$\begin{aligned} \text{Maximize } p &= 10A + 5B \\ &3A + 1B \le 10 \\ &2A + 3B \le 12 \\ &A, B \ge 0 \end{aligned}$$

PA-5. Solve by the simplex method.

$$\begin{aligned} \text{Maximize } p &= 10x_1 + 5x_2 + 1x_3 \\ &1x_1 + 1x_2 + 1x_3 \le 10 \\ &2x_1 + 1x_2 + 4x_3 \le 12 \\ &x_1, x_2, x_3 \ge 0 \end{aligned}$$

PA-6. Solve, using a standard LP computer code.

$$\begin{aligned} \text{Maximize } p &= 11x_1 + 14x_2 + 19x_3 + 14x_4 + 10x_5 \\ &8x_1 + 9x_2 + 14x_3 + 8x_4 + 10x_5 \le 30 \\ &7x_1 + 14x_2 + 12x_3 + 16x_4 + 6x_5 \le 42 \\ &x_1, x_2, x_3, x_4, x_5 \ge 0 \end{aligned}$$

PA-7. A manufacturer of three products, A, B, and C, must decide on how much of each to produce. The marketing department has assured him that whatever he makes they can sell. Limiting his output are three production departments, including fabrication (*F*), assembly (*S*), and painting (*P*). In particular, each department has 1,000 man-hours of capacity available.

In the manufacture of product A, 2 units of *F*, 3 units of *S*, and 1 unit of *P* are consumed. For B, 3 units of *F*, 1 unit of *S*, and 1 unit of *P*. For C, 4 units of *F*, 3 units of *S*, and 4 units of *P*. Assume that products A and C generate \$10 profit each, while B generates \$20. How much of each should be produced?

PA-8. Bill Hold, production manager for the Allied Production Company, must schedule his production for the next two periods. Twenty-two units have been sold for delivery in period 1 and 40 have been sold for delivery in period 2. Regular production capacity is 25 units per time period, while overtime capacity is 10 units per time period. Overtime costs are \$2 per unit in excess of regular production costs. Units can also be produced in one period and sold in the next. The storage cost for this strategy is \$1 per unit per period. Set this up as a linear programming problem.

Hint: Let X_{ij} = amount produced in period i for sale in period j on regular time
Let X_{ioj} = amount produced in period i for sale in period j on overtime

PA-9. The manager of marketing for the World Wide Frisbee Company is once again preparing his annual advertising strategy. His objective is to minimize the total advertising expenditure while achieving at least some minimum level of exposure in each of four audiences. These audiences include teenage, college, singles between 25 and 35, and newlyweds. To reach these audiences, he can use media which include radio, TV, newspapers, magazines, and billboards.

Let X_j = expenditure on medium j
E_i = minimum exposure desired for audience i
e_{ij} = effectiveness of jth advertising medium on ith audience

Write an LP model for this problem. Do the assumptions of LP raise any doubts about this model's relevance?

PA-10. The G-M Bicycle Company produces two types of bicycles. Each one requires four operations. With the recent emphasis on physical fitness, management feels that it can sell as many as are produced. If the profit from bicycle A is $20 per unit and that from B is $35, how many of each should the company produce? The maximum number of bikes which can be processed through each department for either bike A *or* bike B is

	Bike A	*Bike B*
Stamping	1,200	1,400
Assembly	1,000	900
Paint	800	1,200
Inspection	1,000	700

Solve: (*a*) graphically; (*b*) by the simplex method.

appendix B

Tables

Table B-1 Random numbers

53076	98356	71012	72913	57081	50378	24782	59604	68503	87115
67675	66328	31868	81477	44108	30976	97286	11185	85146	80501
11682	77634	35669	48952	11988	76536	47230	5101	56004	447
80779	52122	53345	65387	98605	43790	96415	48056	23627	15127
18002	40596	31530	47181	56785	91219	98542	55551	52010	42465
74217	49876	28094	41249	78516	83703	49785	37480	43103	15747
10332	39877	18937	66983	80203	5878	25213	86229	53043	65949
81641	6372	15397	58928	98026	56147	41803	27564	16801	49977
43916	21062	73384	23830	15617	27292	47003	80081	60913	65541
76522	33375	75162	85546	76477	51924	56899	87600	11637	32607
89274	71515	72226	76519	30405	13914	88805	6010	56571	62411
32280	90189	77336	40196	34756	49649	3071	44055	58885	55252
43878	83500	48961	20485	1404	55651	50495	41115	45496	87123
92040	11812	5970	16516	68018	38973	22272	71589	45584	73886
6359	2591	68104	91446	22993	21631	36380	53439	58588	61660
86313	36575	46153	37869	28847	5297	49395	51115	55731	29636
87662	73748	84288	29618	52395	70084	53295	15741	99974	74435
24285	66547	48913	73717	53551	65409	21248	56997	85000	35454
69098	76804	13315	96064	99615	35505	85816	44036	39764	6913
19215	42140	60830	83484	32894	51061	1645	31020	41339	67819
98467	1399	87220	35500	81485	9378	93859	14486	55827	24842
52206	36683	53151	73242	83242	19358	84254	96160	94802	49007
93275	36216	90302	89493	87642	66725	25401	72950	94381	32159
53993	7594	26512	74082	15915	70083	52465	93485	44128	31367
72734	80239	16993	40338	75399	85135	69427	2434	12818	3234
20823	36035	11101	1463	50068	5666	15313	75499	19621	44791
18203	39387	32534	41814	38175	88661	63836	62019	61531	77848
32030	42255	75125	48994	53716	47878	47411	84658	96763	92145
32626	32699	4949	5213	66642	15814	72257	7122	58807	78031
42589	5796	44046	50050	99642	28893	33291	92302	72065	16855
45818	6379	22194	95072	16463	4221	83866	11458	55554	54878
15188	51600	35924	1054	45465	84325	66904	2428	6880	18112
56103	99032	40717	51157	97265	49433	88443	47471	44185	87971
89959	49977	27971	19762	84439	79377	62707	42935	49401	56964
79082	70637	1453	40502	38065	52271	787	80762	35583	63603
22815	10297	4490	50398	44966	31144	64458	78696	88324	29508
62014	56652	42726	41713	37607	42638	54659	67205	259	57143
68594	76822	31362	80684	58100	28842	83207	58900	70070	40103
77866	66073	79150	37985	43637	42600	17149	95556	96339	72246
45063	57832	12466	54501	11314	96443	75026	50885	27642	93890
12734	19579	3176	47576	48602	45393	85097	31695	9856	67375
64837	54064	78104	1776	60400	68663	45096	90193	81828	92087
56683	73431	70175	43790	96677	58275	50851	93630	88019	27138
7263	98095	12841	26402	65134	94661	9645	58195	71850	4143
48058	25821	89441	36557	28807	6601	42032	54674	82623	80138
48051	19019	48678	40788	20914	17274	18822	53466	84952	87683
25867	34598	86636	57225	10061	94130	83161	13024	7262	97539
70560	25520	90985	66630	30384	61363	11016	17455	98874	84514
11130	30614	38973	22382	80797	53708	25215	88232	38211	67251
70042	12020	12248	37901	60797	46160	44691	89594	88327	32170

From Robert C. Meier, William T. Newell, and Harold L. Pazer, *Simulation in Business and Economics*, © 1969, pp. 333–334. Reprinted by permission of Prentice-Hall, Inc., Englewood Cliffs, N.J.

Table B-1 Random numbers (Continued)

81220	89905	96270	3623	90920	91528	5124	78102	99305	12198
2014	96843	72247	97374	98257	89117	15199	63100	32723	29380
73567	4964	20262	80530	5760	16541	92486	53998	12614	1070
8847	68279	65016	31205	24808	61308	56989	76232	46886	64975
84935	70822	84952	87846	56112	91074	54363	74277	8944	64363
7043	80474	49962	12441	29898	63879	24339	20601	16021	76895
29407	42824	38890	40494	29968	3065	38387	41567	93671	28325
99049	58423	97500	22786	81410	70815	77952	50940	82326	85601
77766	66495	97116	42433	51279	30837	60086	45643	32648	54973
17725	66003	9927	38545	98164	78503	96499	30562	87855	64955
81120	90401	87452	65591	844	70565	30688	12251	40816	49486
37314	78605	98343	58618	90702	41387	15051	15600	60412	68847
85817	44878	74780	7073	10142	27440	94016	70152	20730	44038
50924	66059	64853	69569	43822	42330	49230	87444	57565	47094
27781	31945	57841	20988	99685	70551	16834	82722	78239	35710
88139	46107	92977	40399	35679	89097	96061	96834	62644	80971
58373	47705	76396	8519	42943	42827	42016	38177	33948	43368
57099	85151	85064	99127	35661	78037	34949	35350	32095	6267
40950	81835	99200	8138	65345	10632	36604	75115	39045	93913
57196	81887	50162	10946	48234	68378	62644	80983	54443	53328
346	45443	34641	29396	31907	48096	63436	65724	32923	27460
20551	66428	30398	24441	21103	13002	85275	7796	26772	31961
13237	18786	17842	81901	64066	73767	4053	17150	96371	4372
89865	56546	37401	64846	62861	33533	31696	11715	9585	99254
95711	50585	29939	70075	44808	61069	19937	57696	77096	2724
5360	11764	58574	46894	72658	290	87552	64119	42070	92144
4992	47278	52559	86678	98174	14897	63431	60280	38274	29566
90597	81910	73621	59393	58848	615	9830	42081	3249	20599
19083	11456	53840	56265	58926	14049	22596	92848	12921	5625
96383	15958	15256	19494	19480	74910	36517	89190	87469	82434
5254	7036	73613	50525	70951	92644	10981	82883	37911	70583
13242	23236	26973	31103	23110	48336	1108	98940	50512	58286
2078	59372	38074	32130	41471	61442	89835	27125	81561	27098
98302	17526	69175	52521	49220	54538	47889	58361	36290	63880
77196	1375	63116	48593	56607	5183	37144	9873	84918	53938
98134	51069	10277	84679	17500	53137	59070	38531	84397	38345
43426	36102	77651	52946	69670	101	116	15814	72050	1696
43107	19419	44704	2231	11132	80806	78995	84566	5640	89349
32612	18515	49308	64825	42509	45466	57749	29870	1691	76404
26610	70580	45769	57831	10834	17156	41694	18903	33120	22320
37474	37018	85106	41007	38652	19310	36791	60364	21221	30898
4539	99115	23110	2108	89143	20156	15174	38180	36639	10213
41476	2827	2481	59144	11743	21120	30667	91562	38376	31092
38167	24378	58748	20093	12822	12344	33373	72754	134	33442
7121	57739	20073	93123	85809	41746	70687	51690	25597	66914
37304	68971	50934	76143	57872	12168	58969	38353	8263	89050
51342	79992	72212	62334	73850	48866	26419	81749	13486	64870
85490	21180	89743	36291	65229	87035	51856	90132	21086	96808
42013	35035	19893	14783	50447	37474	37616	77744	44506	5727
93459	18359	94140	92910	74259	76318	31448	65133	47542	14312

Table B-2 Present value of $1.00

Years Hence	1%	2%	4%	6%	8%	10%	12%	14%	15%	16%	18%	20%	22%	24%	25%	26%	28%	30%	35%	40%	45%	50%
1	0.990	0.980	0.962	0.943	0.926	0.909	0.893	0.877	0.870	0.862	0.847	0.833	0.820	0.806	0.800	0.794	0.781	0.769	0.741	0.714	0.690	0.667
2	0.980	0.961	0.925	0.890	0.857	0.826	0.797	0.769	0.756	0.743	0.718	0.694	0.672	0.650	0.640	0.630	0.610	0.592	0.549	0.510	0.476	0.444
3	0.971	0.942	0.889	0.840	0.794	0.751	0.712	0.675	0.658	0.641	0.609	0.579	0.551	0.524	0.512	0.500	0.477	0.455	0.406	0.364	0.328	0.296
4	0.961	0.924	0.855	0.792	0.735	0.683	0.636	0.592	0.572	0.552	0.516	0.482	0.451	0.423	0.410	0.397	0.373	0.350	0.301	0.260	0.226	0.198
5	0.951	0.906	0.822	0.747	0.681	0.621	0.567	0.519	0.497	0.476	0.437	0.402	0.370	0.341	0.328	0.315	0.291	0.269	0.223	0.186	0.156	0.132
6	0.942	0.888	0.790	0.705	0.630	0.564	0.507	0.456	0.432	0.410	0.370	0.335	0.303	0.275	0.262	0.250	0.227	0.207	0.165	0.133	0.108	0.088
7	0.933	0.871	0.760	0.665	0.583	0.513	0.452	0.400	0.376	0.354	0.314	0.279	0.249	0.222	0.210	0.198	0.178	0.159	0.122	0.095	0.074	0.059
8	0.923	0.853	0.731	0.627	0.540	0.467	0.404	0.351	0.327	0.305	0.266	0.233	0.204	0.179	0.168	0.157	0.139	0.123	0.091	0.068	0.051	0.039
9	0.914	0.837	0.703	0.592	0.500	0.424	0.361	0.308	0.284	0.263	0.225	0.194	0.167	0.144	0.134	0.125	0.108	0.094	0.067	0.048	0.035	0.026
10	0.905	0.820	0.676	0.558	0.463	0.386	0.322	0.270	0.247	0.227	0.191	0.162	0.137	0.116	0.107	0.099	0.085	0.073	0.050	0.035	0.024	0.017
11	0.896	0.804	0.650	0.527	0.429	0.350	0.287	0.237	0.215	0.195	0.162	0.135	0.112	0.094	0.086	0.079	0.066	0.056	0.037	0.025	0.017	0.012
12	0.887	0.788	0.625	0.497	0.397	0.319	0.257	0.208	0.187	0.168	0.137	0.112	0.092	0.076	0.069	0.062	0.052	0.043	0.027	0.018	0.012	0.008
13	0.879	0.773	0.601	0.469	0.368	0.290	0.229	0.182	0.163	0.145	0.116	0.093	0.075	0.061	0.055	0.050	0.040	0.033	0.020	0.013	0.008	0.005
14	0.870	0.758	0.577	0.442	0.340	0.263	0.205	0.160	0.141	0.125	0.099	0.078	0.062	0.049	0.044	0.039	0.032	0.025	0.015	0.009	0.006	0.003
15	0.861	0.743	0.555	0.417	0.315	0.239	0.183	0.140	0.123	0.108	0.084	0.065	0.051	0.040	0.035	0.031	0.025	0.020	0.011	0.006	0.004	0.002
16	0.853	0.728	0.534	0.394	0.292	0.218	0.163	0.123	0.107	0.093	0.071	0.054	0.042	0.032	0.028	0.025	0.019	0.015	0.008	0.005	0.003	0.002
17	0.844	0.714	0.513	0.371	0.270	0.198	0.146	0.108	0.093	0.080	0.060	0.045	0.034	0.026	0.023	0.020	0.015	0.012	0.006	0.003	0.002	0.001
18	0.836	0.700	0.494	0.350	0.250	0.180	0.130	0.095	0.081	0.069	0.051	0.038	0.028	0.021	0.018	0.016	0.012	0.009	0.005	0.002	0.001	0.001
19	0.828	0.686	0.475	0.331	0.232	0.164	0.116	0.083	0.070	0.060	0.043	0.031	0.023	0.017	0.014	0.012	0.009	0.007	0.003	0.002	0.001	
20	0.820	0.673	0.456	0.312	0.215	0.149	0.104	0.073	0.061	0.051	0.037	0.026	0.019	0.014	0.012	0.010	0.007	0.005	0.002	0.001	0.001	
21	0.811	0.660	0.439	0.294	0.199	0.135	0.093	0.064	0.053	0.044	0.031	0.022	0.015	0.011	0.009	0.008	0.006	0.004	0.002	0.001		
22	0.803	0.647	0.422	0.278	0.184	0.123	0.083	0.056	0.046	0.038	0.026	0.018	0.013	0.009	0.007	0.006	0.004	0.003	0.001	0.001		
23	0.795	0.634	0.406	0.262	0.170	0.112	0.074	0.049	0.040	0.033	0.022	0.015	0.010	0.007	0.006	0.005	0.003	0.002	0.001			
24	0.788	0.622	0.390	0.247	0.158	0.102	0.066	0.043	0.035	0.028	0.019	0.013	0.008	0.006	0.005	0.004	0.003	0.002	0.001			
25	0.780	0.610	0.375	0.233	0.146	0.092	0.059	0.038	0.030	0.024	0.016	0.010	0.007	0.005	0.004	0.003	0.002	0.001	0.001			
26	0.772	0.598	0.361	0.220	0.135	0.084	0.053	0.033	0.026	0.021	0.014	0.009	0.006	0.004	0.003	0.002	0.002	0.001				
27	0.764	0.586	0.347	0.207	0.125	0.076	0.047	0.029	0.023	0.018	0.011	0.007	0.005	0.003	0.002	0.002	0.001	0.001				
28	0.757	0.574	0.333	0.196	0.116	0.069	0.042	0.026	0.020	0.016	0.010	0.006	0.004	0.002	0.002	0.002	0.001	0.001				
29	0.749	0.563	0.321	0.185	0.107	0.063	0.037	0.022	0.017	0.014	0.008	0.005	0.003	0.002	0.002	0.001	0.001	0.001				
30	0.742	0.552	0.308	0.174	0.099	0.057	0.033	0.020	0.015	0.012	0.007	0.004	0.003	0.002	0.001	0.001	0.001					
40	0.672	0.453	0.208	0.097	0.046	0.022	0.011	0.005	0.004	0.003	0.001	0.001										
50	0.608	0.372	0.141	0.054	0.021	0.009	0.003	0.001	0.001	0.001												

Reproduced with permission from Robert N. Anthony, *Management Accounting Principles*, rev. ed., Richard D. Irwin, Inc., Homewood, Ill., 1970.

Table B-3 Present value of annuities: $1 per year

Years (N)	1%	2%	4%	6%	8%	10%	12%	14%	15%	16%	18%	20%	22%	24%	25%	26%	28%	30%	35%	40%	45%	50%
1	0.990	0.980	0.962	0.943	0.926	0.909	0.893	0.877	0.870	0.862	0.847	0.833	0.820	0.806	0.800	0.794	0.781	0.769	0.741	0.714	0.690	0.667
2	1.970	1.942	1.886	1.833	1.783	1.736	1.690	1.647	1.626	1.605	1.566	1.528	1.492	1.457	1.440	1.424	1.392	1.361	1.289	1.224	1.165	1.111
3	2.941	2.884	2.775	2.673	2.577	2.487	2.402	2.322	2.283	2.246	2.174	2.106	2.042	1.981	1.952	1.923	1.868	1.816	1.696	1.589	1.493	1.407
4	3.902	3.808	3.630	3.465	3.312	3.170	3.037	2.914	2.855	2.798	2.690	2.589	2.494	2.404	2.362	2.320	2.241	2.166	1.997	1.849	1.720	1.605
5	4.853	4.713	4.452	4.212	3.993	3.791	3.605	3.433	3.352	3.274	3.127	2.991	2.864	2.745	2.689	2.635	2.532	2.436	2.220	2.035	1.876	1.737
6	5.795	5.601	5.242	4.917	4.623	4.355	4.111	3.889	3.784	3.685	3.498	3.326	3.167	3.020	2.951	2.885	2.759	2.643	2.385	2.168	1.983	1.824
7	6.728	6.472	6.002	5.582	5.206	4.868	4.564	4.288	4.160	4.039	3.812	3.605	3.416	3.242	3.161	3.083	2.937	2.802	2.508	2.263	2.057	1.883
8	7.652	7.325	6.733	6.210	5.747	5.335	4.968	4.639	4.487	4.344	4.078	3.837	3.619	3.421	3.329	3.241	3.076	2.925	2.598	2.331	2.108	1.922
9	8.566	8.162	7.435	6.802	6.247	5.759	5.328	4.946	4.772	4.607	4.303	4.031	3.786	3.566	3.463	3.366	3.184	3.019	2.665	2.379	2.144	1.948
10	9.471	8.983	8.111	7.360	6.710	6.145	5.650	5.216	5.019	4.833	4.494	4.192	3.923	3.682	3.571	3.465	3.269	3.092	2.715	2.414	2.168	1.965
11	10.368	9.787	8.760	7.887	7.139	6.495	5.937	5.453	5.234	5.029	4.656	4.327	4.035	3.776	3.656	3.544	3.335	3.147	2.757	2.438	2.185	1.977
12	11.255	10.575	9.385	8.384	7.536	6.814	6.194	5.660	5.421	5.197	4.793	4.439	4.127	3.851	3.725	3.606	3.387	3.190	2.779	2.456	2.196	1.985
13	12.134	11.343	9.986	8.853	7.904	7.103	6.424	5.842	5.583	5.342	4.910	4.533	4.203	3.912	3.780	3.656	3.427	3.223	2.799	2.468	2.204	1.990
14	13.004	12.106	10.563	9.295	8.244	7.367	6.628	6.002	5.724	5.468	5.008	4.611	4.265	3.962	3.824	3.695	3.459	3.249	2.814	2.477	2.210	1.993
15	13.865	12.849	11.118	9.712	8.559	7.606	6.811	6.142	5.847	5.575	5.092	4.675	4.315	4.001	3.859	3.726	3.483	3.268	2.825	2.484	2.214	1.995
16	14.718	13.578	11.652	10.106	8.851	7.824	6.974	6.265	5.954	5.669	5.162	4.730	4.357	4.033	3.887	3.751	3.503	3.283	2.834	2.489	2.216	1.997
17	15.562	14.292	12.166	10.477	9.122	8.022	7.120	6.373	6.047	5.749	5.222	4.775	4.391	4.059	3.910	3.771	3.518	3.295	2.840	2.492	2.218	1.998
18	16.398	14.992	12.659	10.828	9.372	8.201	7.250	6.467	6.128	5.818	5.273	4.812	4.419	4.080	3.928	3.786	3.529	3.304	2.844	2.494	2.219	1.999
19	17.226	15.678	13.134	11.158	9.604	8.365	7.366	6.550	6.198	5.877	5.316	4.844	4.442	4.097	3.942	3.799	3.539	3.311	2.848	2.496	2.220	1.999
20	18.046	16.351	13.590	11.470	9.818	8.514	7.469	6.623	6.259	5.929	5.353	4.870	4.460	4.110	3.954	3.808	3.546	3.316	2.850	2.497	2.221	1.999
21	18.857	17.011	14.029	11.764	10.017	8.649	7.562	6.687	6.312	5.973	5.384	4.891	4.476	4.121	3.963	3.816	3.551	3.320	2.852	2.498	2.221	2.000
22	19.660	17.658	14.451	12.042	10.201	8.772	7.645	6.743	6.359	6.011	5.410	4.909	4.488	4.130	3.970	3.822	3.556	3.323	2.853	2.498	2.222	2.000
23	20.456	18.292	14.857	12.303	10.371	8.883	7.718	6.792	6.399	6.044	5.432	4.925	4.499	4.137	3.976	3.827	3.559	3.325	2.854	2.499	2.222	2.000
24	21.243	18.914	15.247	12.550	10.529	8.985	7.784	6.835	6.434	6.073	5.451	4.937	4.507	4.143	3.981	3.831	3.562	3.327	2.855	2.499	2.222	2.000
25	22.023	19.523	15.622	12.783	10.675	9.077	7.843	6.873	6.464	6.097	5.467	4.948	4.514	4.147	3.985	3.834	3.564	3.329	2.856	2.499	2.222	2.000
26	22.795	20.121	15.983	13.003	10.810	9.161	7.896	6.906	6.491	6.118	5.480	4.956	4.520	4.151	3.988	3.837	3.566	3.330	2.856	2.500	2.222	2.000
27	23.560	20.707	16.330	13.211	10.935	9.237	7.943	6.935	6.514	6.136	5.492	4.964	4.524	4.154	3.990	3.839	3.567	3.331	2.856	2.500	2.222	2.000
28	24.316	21.281	16.663	13.406	11.051	9.307	7.984	6.961	6.534	6.152	5.502	4.970	4.528	4.157	3.992	3.840	3.568	3.331	2.857	2.500	2.222	2.000
29	25.066	21.844	16.984	13.591	11.158	9.370	8.022	6.983	6.551	6.166	5.510	4.975	4.531	4.159	3.994	3.841	3.569	3.332	2.857	2.500	2.222	2.000
30	25.808	22.396	17.292	13.765	11.258	9.427	8.055	7.003	6.566	6.177	5.517	4.979	4.534	4.160	3.995	3.842	3.569	3.332	2.857	2.500	2.222	2.000
40	32.835	27.355	19.793	15.046	11.925	9.779	8.244	7.105	6.642	6.234	5.548	4.997	4.544	4.166	3.999	3.846	3.571	3.333	2.857	2.500	2.222	2.000
50	39.196	31.424	21.482	15.762	12.234	9.915	8.304	7.133	6.661	6.246	5.554	4.999	4.545	4.167	4.000	3.846	3.571	3.333	2.857	2.500	2.222	2.000

Reproduced with permission from Robert N. Anthony, *Management Accounting Principles*, rev. ed., Richard D. Irwin, Inc., Homewood, Ill., 1970.

Table B-4 Capital-recovery factors for interest rates from 6% to 50%

$$\frac{i(1+i)^n}{(1+i)^n - 1}$$

RP i-n
(xxxxx)

n	i										n
	6%	8%	10%	12%	15%	20%	25%	30%	40%	50%	
1	1.06000	1.08000	1.10000	1.12000	1.15000	1.20000	1.25000	1.30000	1.40000	1.50000	1
2	0.54544	0.56077	0.57619	0.59170	0.61512	0.65455	0.69444	0.73478	0.81667	0.90000	2
3	0.37411	0.38803	0.40211	0.41635	0.43798	0.47473	0.51230	0.55063	0.62936	0.71053	3
4	0.28859	0.30192	0.31547	0.32923	0.35027	0.38629	0.42344	0.46163	0.54077	0.62308	4
5	0.23740	0.25046	0.26380	0.27741	0.29832	0.33438	0.37184	0.41058	0.49136	0.57582	5
6	0.20336	0.21632	0.22961	0.24323	0.26424	0.30071	0.33882	0.37840	0.46126	0.54812	6
7	0.17914	0.19207	0.20541	0.21912	0.24036	0.27742	0.31634	0.35687	9.44192	0.53108	7
8	0.16104	0.17401	0.18744	0.20130	0.22285	0.26061	0.30040	0.34191	0.42804	0.52030	8
9	0.14702	0.16008	0.17364	0.18768	0.20957	0.24808	0.28876	0.33123	0.42034	0.51335	9
10	0.13587	0.14903	0.16275	0.17698	0.19925	0.23852	0.28007	0.32346	0.41432	0.50823	10
11	0.12679	0.14008	0.15396	0.16842	0.19107	0.23110	0.27349	0.31773	0.41013	0.50585	11
12	0.11928	0.13270	0.14676	0.16144	0.18448	0.22526	0.26845	0.31345	0.40718	0.50388	12
13	0.11296	0.12652	0.14078	0.15568	0.17911	0.22062	0.26454	0.31024	0.40510	0.50258	13
14	0.10758	0.12130	0.13575	0.15087	0.17469	0.21689	0.26150	0.30782	0.40363	0.50172	14
15	0.10296	0.11683	0.13147	0.14682	0.17102	0.21388	0.25912	0.30598	0.40259	0.50114	15
16	0.09895	0.11298	0.12782	0.14339	0.16795	0.21144	0.25724	0.30458	0.40185	0.50076	16
17	0.09544	0.10963	0.12466	0.14046	0.16537	0.20944	0.25576	0.30351	0.40132	0.50051	17
18	0.09236	0.10670	0.12193	0.13794	0.16319	0.20781	0.25459	0.30269	0.40094	0.50034	18
19	0.08962	0.10413	0.11955	0.13576	0.16134	0.20646	0.25366	0.30206	0.40067	0.50023	19
20	0.08718	0.10185	0.11746	0.13388	0.15976	0.20536	0.25292	0.30159	0.40048	0.50016	20
25	0.07823	0.09368	0.11017	0.12750	0.15470	0.20212	0.25095	0.30043	0.40009	0.50002	25
30	0.07265	0.08883	0.10608	0.12414	0.15230	0.20085	0.25031	0.30011	0.40002	0.50000	30
40	0.06646	0.08386	0.10226	0.12130	0.15056	0.20014	0.25003	0.30008	0.40001	0.50000	40
50	0.06344	0.08174	0.10086	0.12042	0.15014	0.20002	0.25000	0.30001	0.40000	0.50000	50
100	0.06018	0.08004	0.10001	0.12000	0.15000	0.20000	0.25000	0.30000	0.40000	0.50000	100
∞	0.06000	0.08000	0.10000	0.12000	0.15000	0.20000	0.25000	0.30000	0.40000	0.50000	∞

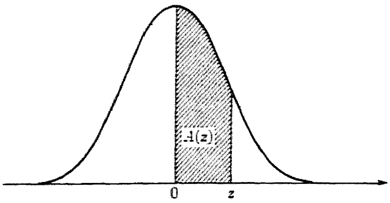
H. G. Thuesen, *Engineering Economy*, 2d ed., © 1957, p. 569. Reprinted by permission of Prentice-Hall, Inc., Englewood Cliffs, N.J.

Table B-5 Normal curve areas

Area under the standard normal curve from 0 to z, shown shaded, is $A(z)$.

Examples. If Z is the standard normal random variable and $z = 1.54$, then

$$\begin{aligned} A(z) = P(0 < Z < z) &= .4382 \\ P(Z > z) &= .0618 \\ P(Z < z) &= .9382 \\ P(|Z| < z) &= .8764 \end{aligned}$$



z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.0000	.0040	.0080	.0120	.0160	.0199	.0239	.0279	.0319	.0359
0.1	.0398	.0438	.0478	.0517	.0557	.0596	.0636	.0675	.0714	.0753
0.2	.0793	.0832	.0871	.0910	.0948	.0987	.1026	.1064	.1103	.1141
0.3	.1179	.1217	.1255	.1293	.1331	.1368	.1406	.1443	.1480	.1517
0.4	.1554	.1591	.1628	.1664	.1700	.1736	.1772	.1808	.1844	.1879
0.5	.1915	.1950	.1985	.2019	.2054	.2088	.2123	.2157	.2190	.2224
0.6	.2257	.2291	.2324	.2357	.2389	.2422	.2454	.2486	.2517	.2549
0.7	.2580	.2611	.2642	.2673	.2704	.2734	.2764	.2794	.2823	.2852
0.8	.2881	.2910	.2939	.2967	.2995	.3023	.3051	.3078	.3106	.3133
0.9	.3159	.3186	.3212	.3238	.3264	.3289	.3315	.3340	.3365	.3389
1.0	.3413	.3438	.3461	.3485	.3508	.3531	.3554	.3577	.3599	.3621
1.1	.3643	.3665	.3686	.3708	.3729	.3749	.3770	.3790	.3810	.3830
1.2	.3849	.3869	.3888	.3907	.3925	.3944	.3962	.3980	.3997	.4015
1.3	.4032	.4049	.4066	.4082	.4099	.4115	.4131	.4147	.4162	.4177
1.4	.4192	.4207	.4222	.4236	.4251	.4265	.4279	.4292	.4306	.4319
1.5	.4332	.4345	.4357	.4370	.4382	.4394	.4406	.4418	.4429	.4441
1.6	.4452	.4463	.4474	.4484	.4495	.4505	.4515	.4525	.4535	.4545
1.7	.4554	.4564	.4573	.4582	.4591	.4599	.4608	.4616	.4625	.4633
1.8	.4641	.4649	.4656	.4664	.4671	.4678	.4686	.4693	.4699	.4706
1.9	.4713	.4719	.4726	.4732	.4738	.4744	.4750	.4756	.4761	.4767
2.0	.4772	.4778	.4783	.4788	.4793	.4798	.4803	.4808	.4812	.4817
2.1	.4821	.4826	.4830	.4834	.4838	.4842	.4846	.4850	.4854	.4857
2.2	.4861	.4864	.4868	.4871	.4875	.4878	.4881	.4884	.4887	.4890
2.3	.4893	.4896	.4898	.4901	.4904	.4906	.4909	.4911	.4913	.4916
2.4	.4918	.4920	.4922	.4925	.4927	.4929	.4931	.4932	.4934	.4936
2.5	.4938	.4940	.4941	.4943	.4945	.4946	.4948	.4949	.4951	.4952
2.6	.4953	.4955	.4956	.4957	.4959	.4960	.4961	.4962	.4963	.4964
2.7	.4965	.4966	.4967	.4968	.4969	.4970	.4971	.4972	.4973	.4974
2.8	.4974	.4975	.4976	.4977	.4977	.4978	.4979	.4979	.4980	.4981
2.9	.4981	.4982	.4982	.4983	.4984	.4984	.4985	.4985	.4986	.4986
3.0	.4987	.4987	.4987	.4988	.4988	.4989	.4989	.4989	.4990	.4990

From Frederick Mosteller, Robert E. K. Rourke, and George B. Thomas, Jr., *Probability and Statistics,* Addison-Wesley Publishing Co., Reading, Mass., 1961, p. 368.

Table B-6 Summation of terms of the Poisson formula

$$1{,}000 \times P(r \leq c;\mu) = 1{,}000 \sum_{r=0}^{r=c} \frac{u^r e^{-\mu}}{r!}$$

1,000 × probability of *c* or less occurrences of event that has average number of occurrences equal to *c′* or *np′*

c′ or *np′* \ *c*	0	1	2	3	4	5	6	7	8	9
0.02	980	1,000								
0.04	961	999	1,000							
0.06	942	998	1,000							
0.08	923	997	1,000							
0.10	905	995	1,000							
0.15	861	990	999	1,000						
0.20	819	982	999	1,000						
0.25	779	974	998	1,000						
0.30	741	963	996	1,000						
0.35	705	951	994	1,000						
0.40	670	938	992	999	1,000					
0.45	638	925	989	999	1,000					
0.50	607	910	986	998	1,000					
0.55	577	894	982	998	1,000					
0.60	549	878	977	997	1,000					
0.65	522	861	972	996	999	1,000				
0.70	497	844	966	994	999	1,000				
0.75	472	827	959	993	999	1,000				
0.80	449	809	953	991	999	1,000				
0.85	427	791	945	989	998	1,000				
0.90	407	772	937	987	998	1,000				
0.95	387	754	929	984	997	1,000				
1.00	368	736	920	981	996	999	1,000			
1.1	333	699	900	974	995	999	1,000			
1.2	301	663	879	966	992	998	1,000			
1.3	273	627	857	957	989	998	1,000			
1.4	247	592	833	946	986	997	999	1,000		
1.5	223	558	809	934	981	996	999	1,000		
1.6	202	525	783	921	976	994	999	1,000		
1.7	183	493	757	907	970	992	998	1,000		
1.8	165	463	731	891	964	990	997	999	1,000	
1.9	150	434	704	875	956	987	997	999	1,000	
2.0	135	406	677	857	947	983	995	999	1,000	

.92

From *Statistical Quality Control* by E. L. Grant. Copyright, 1964, by McGraw-Hill, Inc. Used by permission of McGraw-Hill Book Co., pp. 567–569.

Table B-6 Summation of terms of the Poisson formula (Continued)

c c' or np'	0	1	2	3	4	5	6	7	8	9
2.2	111	355	623	819	928	975	993	998	1,000	
2.4	091	308	570	779	904	964	988	997	999	1,000
2.6	074	267	518	736	877	951	983	995	999	1,000
2.8	061	231	469	692	848	935	976	992	998	999
3.0	050	199	423	647	815	916	966	988	996	999
3.2	041	171	380	603	781	895	955	983	994	998
3.4	033	147	340	558	744	871	942	977	992	997
3.6	027	126	303	515	706	844	927	969	988	996
3.8	022	107	269	473	668	816	909	960	984	994
4.0	018	092	238	433	629	785	889	949	979	992
62.27 → 4.2	015	078	210	395	590	753	867	936	972	989
4.4	012	066	185	359	551	720	844	921	964	985
4.6	010	056	163	326	513	686	818	905	955	980
4.8	008	048	143	294	476	651	791	887	944	975
5.0	007	040	125	265	440	616	762	867	932	968
5.2	006	034	109	238	406	581	732	845	918	960
5.4	005	029	095	213	373	546	702	822	903	951
5.6	004	024	082	191	342	512	670	797	886	941
5.8	003	021	072	170	313	478	638	771	867	929
6.0	002	017	062	151	285	446	606	744	847	916

	10	11	12	13	14	15	16
2.8	1,000						
3.0	1,000						
3.2	1,000						
3.4	999	1,000					
3.6	999	1,000					
3.8	998	999	1,000				
4.0	997	999	1,000				
4.2	996	999	1,000				
4.4	994	998	999	1,000			
4.6	992	997	999	1,000			
4.8	990	996	999	1,000			
5.0	986	995	998	999	1,000		
5.2	982	993	997	999	1,000		
5.4	977	990	996	999	1,000		
5.6	972	988	995	998	999	1,000	
5.8	965	984	993	997	999	1,000	
6.0	957	980	991	996	999	999	1,000

Table B-6 Summation of terms of the Poisson formula (Continued)

c' or np' \ c	0	1	2	3	4	5	6	7	8	9
6.2	002	015	054	134	259	414	574	716	826	902
6.4	002	012	046	119	235	384	542	687	803	886
6.6	001	010	040	105	213	355	511	658	780	869
6.8	001	009	034	093	192	327	480	628	755	850
7.0	001	007	030	082	173	301	450	599	729	830
7.2	001	006	025	072	156	276	420	569	703	810
7.4	001	005	022	063	140	253	392	539	676	788
7.6	001	004	019	055	125	231	365	510	648	765
7.8	000	004	016	048	112	210	338	481	620	741
8.0	000	003	014	042	100	191	313	453	593	717
8.5	000	002	009	030	074	150	256	386	523	653
9.0	000	001	006	021	055	116	207	324	456	587
9.5	000	001	004	015	040	089	165	269	392	522
10.0	000	000	003	010	029	067	130	220	333	458

	10	11	12	13	14	15	16	17	18	19
6.2	949	975	989	995	998	999	1,000			
6.4	939	969	986	994	997	999	1,000			
6.6	927	963	982	992	997	999	999	1,000		
6.8	915	955	978	990	996	998	999	1,000		
7.0	901	947	973	987	994	998	999	1,000		
7.2	887	937	967	984	993	997	999	999	1,000	
7.4	871	926	961	980	991	996	998	999	1,000	
7.6	854	915	954	976	989	995	998	999	1,000	
7.8	835	902	945	971	986	993	997	999	1,000	
8.0	816	888	936	966	983	992	996	998	999	1,000
8.5	763	849	909	949	973	986	993	997	999	999
9.0	706	803	876	926	959	978	989	995	998	999
9.5	645	752	836	898	940	967	982	991	996	998
10.0	583	697	792	864	917	951	973	986	993	997

	20	21	22
8.5	1,000		
9.0	1,000		
9.5	999	1,000	
10.0	998	999	1,000

Table B-7 Percentage points of the F distribution

Example: For $\nu_1 = 12$ and $\nu_2 = 14$
$P(F > 2.79) = 0.05$
$P(F > 4.40) = 0.01$

The obtained F is significant at a given level if it is equal to or *greater than* the value shown in the table. 0.05 (light row) and 0.01 (dark row) points for the distribution of F

Degrees of freedom for lesser mean square	Degrees of freedom for greater mean square																							
	1	2	3	4	5	6	7	8	9	10	11	12	14	16	20	24	30	40	50	75	100	200	500	∞
1	161 **4052**	200 **4999**	216 **5403**	225 **5625**	230 **5764**	234 **5859**	237 **5928**	239 **5981**	241 **6022**	242 **6056**	243 **6082**	244 **6106**	245 **6142**	246 **6169**	248 **6208**	249 **6234**	250 **6258**	251 **6286**	252 **6302**	253 **6323**	253 **6334**	254 **6352**	254 **6361**	254 **6366**
2	18.51 **98.49**	19.00 **99.01**	19.16 **99.17**	19.25 **99.25**	19.30 **99.30**	19.33 **99.33**	19.36 **99.34**	19.37 **99.36**	19.38 **99.38**	19.39 **99.40**	19.40 **99.41**	19.41 **99.42**	19.42 **99.43**	19.43 **99.44**	19.44 **99.45**	19.45 **99.46**	19.46 **99.47**	19.47 **99.48**	19.47 **99.48**	19.48 **99.49**	19.49 **99.49**	19.49 **99.49**	19.50 **99.50**	19.50 **99.50**
3	10.13 **34.12**	9.55 **30.81**	9.28 **29.46**	9.12 **28.71**	9.01 **28.24**	8.94 **27.91**	8.88 **27.67**	8.84 **27.49**	8.81 **27.34**	8.78 **27.23**	8.76 **27.13**	8.74 **27.05**	8.71 **26.92**	8.69 **26.83**	8.66 **26.69**	8.64 **26.60**	8.62 **26.50**	8.60 **26.41**	8.58 **26.30**	8.57 **26.27**	8.56 **26.23**	8.54 **26.18**	8.54 **26.14**	8.53 **26.12**
4	7.71 **21.20**	6.94 **18.00**	6.59 **16.69**	6.39 **15.98**	6.26 **15.52**	6.16 **15.21**	6.09 **14.98**	6.04 **14.80**	6.00 **14.66**	5.96 **14.54**	5.93 **14.45**	5.91 **14.37**	5.87 **14.24**	5.84 **14.15**	5.80 **14.02**	5.77 **13.93**	5.74 **13.83**	5.71 **13.74**	5.70 **13.69**	5.68 **13.61**	5.66 **13.57**	5.65 **13.52**	5.64 **13.48**	5.63 **13.46**
5	6.61 **16.26**	5.79 **13.27**	5.41 **12.06**	5.19 **11.39**	5.05 **10.97**	4.95 **10.67**	4.88 **10.45**	4.82 **10.27**	4.78 **10.15**	4.74 **10.05**	4.70 **9.96**	4.68 **9.89**	4.64 **9.77**	4.60 **9.68**	4.56 **9.55**	4.53 **9.47**	4.50 **9.38**	4.46 **9.29**	4.44 **9.24**	4.42 **9.17**	4.40 **9.13**	4.38 **9.07**	4.37 **9.04**	4.36 **9.02**
6	5.99 **13.74**	5.14 **10.92**	4.76 **9.78**	4.53 **9.15**	4.39 **8.75**	4.28 **8.47**	4.21 **8.26**	4.15 **8.10**	4.10 **7.98**	4.06 **7.87**	4.03 **7.79**	4.00 **7.72**	3.96 **7.60**	3.92 **7.52**	3.87 **7.39**	3.84 **7.31**	3.81 **7.23**	3.77 **7.14**	3.75 **7.09**	3.72 **7.02**	3.71 **6.99**	3.69 **6.94**	3.68 **6.90**	3.67 **6.88**
7	5.59 **12.25**	4.74 **9.55**	4.35 **8.45**	4.12 **7.85**	3.97 **7.46**	3.87 **7.19**	3.79 **7.00**	3.73 **6.84**	3.68 **6.71**	3.63 **6.62**	3.60 **6.54**	3.57 **6.47**	3.52 **6.35**	3.49 **6.27**	3.44 **6.15**	3.41 **6.07**	3.38 **5.98**	3.34 **5.90**	3.32 **5.85**	3.29 **5.78**	3.28 **5.75**	3.25 **5.70**	3.24 **5.67**	3.23 **5.65**
8	5.32 **11.26**	4.46 **8.65**	4.07 **7.59**	3.84 **7.01**	3.69 **6.63**	3.58 **6.37**	3.50 **6.19**	3.44 **6.03**	3.39 **5.91**	3.34 **5.82**	3.31 **5.74**	3.28 **5.67**	3.23 **5.56**	3.20 **5.48**	3.15 **5.36**	3.12 **5.28**	3.08 **5.20**	3.05 **5.11**	3.03 **5.06**	3.00 **5.00**	2.98 **4.96**	2.96 **4.91**	2.94 **4.88**	2.93 **4.86**
9	5.12 **10.56**	4.26 **8.02**	3.86 **6.99**	3.63 **6.42**	3.48 **6.06**	3.37 **5.80**	3.29 **5.62**	3.23 **5.47**	3.18 **5.35**	3.13 **5.26**	3.10 **5.18**	3.07 **5.11**	3.02 **5.00**	2.98 **4.92**	2.93 **4.80**	2.90 **4.73**	2.86 **4.64**	2.82 **4.56**	2.80 **4.51**	2.77 **4.45**	2.76 **4.41**	2.73 **4.36**	2.72 **4.33**	2.71 **4.31**
10	4.96 **10.04**	4.10 **7.56**	3.71 **6.55**	3.48 **5.99**	3.33 **5.64**	3.22 **5.39**	3.14 **5.21**	3.07 **5.06**	3.02 **4.95**	2.97 **4.85**	2.94 **4.78**	2.91 **4.71**	2.86 **4.60**	2.82 **4.52**	2.77 **4.41**	2.74 **4.33**	2.70 **4.25**	2.67 **4.17**	2.64 **4.12**	2.61 **4.05**	2.59 **4.01**	2.56 **3.96**	2.55 **3.93**	2.54 **3.91**
11	4.84 **9.65**	3.98 **7.20**	3.59 **6.22**	3.36 **5.67**	3.20 **5.32**	3.09 **5.07**	3.01 **4.88**	2.95 **4.74**	2.90 **4.63**	2.86 **4.54**	2.82 **4.46**	2.79 **4.40**	2.74 **4.29**	2.70 **4.21**	2.65 **4.10**	2.61 **4.02**	2.57 **3.94**	2.53 **3.86**	2.50 **3.80**	2.47 **3.74**	2.45 **3.70**	2.42 **3.66**	2.41 **3.62**	2.40 **3.60**
12	4.75 **9.33**	3.88 **6.93**	3.49 **5.95**	3.26 **5.41**	3.11 **5.06**	3.00 **4.82**	2.92 **4.65**	2.85 **4.50**	2.80 **4.39**	2.76 **4.30**	2.72 **4.22**	2.69 **4.16**	2.64 **4.05**	2.60 **3.98**	2.54 **3.86**	2.50 **3.78**	2.46 **3.70**	2.42 **3.61**	2.40 **3.56**	2.36 **3.49**	2.35 **3.46**	2.32 **3.41**	2.31 **3.38**	2.30 **3.36**
13	4.67 **9.07**	3.80 **6.70**	3.41 **5.74**	3.18 **5.20**	3.02 **4.86**	2.92 **4.62**	2.84 **4.44**	2.77 **4.30**	2.72 **4.19**	2.67 **4.10**	2.63 **4.02**	2.60 **3.96**	2.55 **3.85**	2.51 **3.78**	2.46 **3.67**	2.42 **3.59**	2.38 **3.51**	2.34 **3.42**	2.32 **3.37**	2.28 **3.30**	2.26 **3.27**	2.24 **3.21**	2.22 **3.18**	2.21 **3.16**

From: Ann Hughes and Dennis Grawoig, *Statistics: A Foundation for Analysis*, Addison-Wesley Publishing Co., Reading, Mass., 1971, pp. 501–503.

Table B-7 Percentage points of the F distribution (Continued)

0.05 (light row) and 0.01 (dark row) points for the distribution of F

	Degrees of freedom for greater mean square																						
Degrees of freedom for lesser mean square	1	2	3	4	5	6	7	8	9	10	11	12	14	16	20	24	30	40	50	75	100	200	500	∞
14	4.60 **8.86**	3.74 **6.51**	3.34 **5.56**	3.11 **5.03**	2.96 **4.69**	2.85 **4.46**	2.77 **4.28**	2.70 **4.14**	2.65 **4.03**	2.60 **3.94**	2.56 **3.86**	2.53 **3.80**	2.48 **3.70**	2.44 **3.62**	2.39 **3.51**	2.35 **3.43**	2.31 **3.34**	2.27 **3.26**	2.24 **3.21**	2.21 **3.14**	2.19 **3.11**	2.16 **3.06**	2.14 **3.02**	2.13 **3.00**
15	4.54 **8.68**	3.68 **6.36**	3.29 **5.42**	3.06 **4.89**	2.90 **4.56**	2.79 **4.32**	2.70 **4.14**	2.64 **4.00**	2.59 **3.89**	2.55 **3.80**	2.51 **3.73**	2.48 **3.67**	2.43 **3.56**	2.39 **3.48**	2.33 **3.36**	2.29 **3.29**	2.25 **3.20**	2.21 **3.12**	2.18 **3.07**	2.15 **3.00**	2.12 **2.97**	2.10 **2.92**	2.08 **2.89**	2.07 **2.87**
16	4.49 **8.53**	3.63 **6.23**	3.24 **5.29**	3.01 **4.77**	2.85 **4.44**	2.74 **4.20**	2.66 **4.03**	2.59 **3.89**	2.54 **3.78**	2.49 **3.69**	2.45 **3.61**	2.42 **3.55**	2.37 **3.45**	2.33 **3.37**	2.28 **3.25**	2.24 **3.18**	2.20 **3.10**	2.16 **3.01**	2.13 **2.96**	2.09 **2.89**	2.07 **2.86**	2.04 **2.80**	2.02 **2.77**	2.01 **2.75**
17	4.45 **8.40**	3.59 **6.11**	3.20 **5.18**	2.96 **4.67**	2.81 **4.34**	2.70 **4.10**	2.62 **3.93**	2.55 **3.79**	2.50 **3.68**	2.45 **3.59**	2.41 **3.52**	2.38 **3.45**	2.33 **3.35**	2.29 **3.27**	2.23 **3.16**	2.19 **3.08**	2.15 **3.00**	2.11 **2.92**	2.08 **2.86**	2.04 **2.79**	2.02 **2.76**	1.99 **2.70**	1.97 **2.67**	1.96 **2.65**
18	4.41 **8.28**	3.55 **6.01**	3.16 **5.09**	2.93 **4.58**	2.77 **4.25**	2.66 **4.01**	2.58 **3.85**	2.51 **3.71**	2.46 **3.60**	2.41 **3.51**	2.37 **3.44**	2.34 **3.37**	2.29 **3.27**	2.25 **3.19**	2.19 **3.07**	2.15 **3.00**	2.11 **2.91**	2.07 **2.83**	2.04 **2.78**	2.00 **2.71**	1.98 **2.68**	1.95 **2.62**	1.93 **2.59**	1.92 **2.57**
19	4.38 **8.18**	3.52 **5.93**	3.13 **5.01**	2.90 **4.50**	2.74 **4.17**	2.63 **3.94**	2.55 **3.77**	2.48 **3.63**	2.43 **3.52**	2.38 **3.43**	2.34 **3.36**	2.31 **3.30**	2.26 **3.19**	2.21 **3.12**	2.15 **3.00**	2.11 **2.92**	2.07 **2.84**	2.02 **2.76**	2.00 **2.70**	1.96 **2.63**	1.94 **2.60**	1.91 **2.54**	1.90 **2.51**	1.88 **2.49**
20	4.35 **8.10**	3.49 **5.85**	3.10 **4.94**	2.87 **4.43**	2.71 **4.10**	2.60 **3.87**	2.52 **3.71**	2.45 **3.56**	2.40 **3.45**	2.35 **3.37**	2.31 **3.30**	2.28 **3.23**	2.23 **3.13**	2.18 **3.05**	2.12 **2.94**	2.08 **2.86**	2.04 **2.77**	1.99 **2.69**	1.96 **2.63**	1.92 **2.56**	1.90 **2.53**	1.87 **2.47**	1.85 **2.44**	1.84 **2.42**
21	4.32 **8.02**	3.47 **5.78**	3.07 **4.87**	2.84 **4.37**	2.68 **4.04**	2.57 **3.81**	2.49 **3.65**	2.42 **3.51**	2.37 **3.40**	2.32 **3.31**	2.28 **3.24**	2.25 **3.17**	2.20 **3.07**	2.15 **2.99**	2.09 **2.88**	2.05 **2.80**	2.00 **2.72**	1.96 **2.63**	1.93 **2.58**	1.80 **2.51**	1.87 **2.47**	1.84 **2.42**	1.82 **2.38**	1.81 **2.36**
22	4.30 **7.94**	3.44 **5.72**	3.05 **4.82**	2.82 **4.31**	2.66 **3.99**	2.55 **3.76**	2.47 **3.59**	2.40 **3.45**	2.35 **3.35**	2.30 **3.26**	2.26 **3.18**	2.23 **3.12**	2.18 **3.02**	2.13 **2.94**	2.07 **2.83**	2.03 **2.75**	1.98 **2.67**	1.93 **2.58**	1.91 **2.53**	1.87 **2.46**	1.84 **2.42**	1.81 **2.37**	1.80 **2.33**	1.78 **2.31**
23	4.28 **7.88**	3.42 **5.66**	3.03 **4.76**	2.80 **4.26**	2.64 **3.94**	2.53 **3.71**	2.45 **3.54**	2.38 **3.41**	2.32 **3.30**	2.28 **3.21**	2.24 **3.14**	2.20 **3.07**	2.14 **2.97**	2.10 **2.89**	2.04 **2.78**	2.00 **2.70**	1.96 **2.62**	1.91 **2.53**	1.88 **2.48**	1.84 **2.41**	1.82 **2.37**	1.79 **2.32**	1.77 **2.28**	1.76 **2.26**
24	4.26 **7.82**	3.40 **5.61**	3.01 **4.72**	2.78 **4.22**	2.62 **3.90**	2.51 **3.67**	2.43 **3.50**	2.36 **3.36**	2.30 **3.25**	2.26 **3.17**	2.22 **3.09**	2.18 **3.03**	2.13 **2.93**	2.09 **2.85**	2.02 **2.74**	1.98 **2.66**	1.94 **2.58**	1.89 **2.49**	1.86 **2.44**	1.82 **2.36**	1.80 **2.33**	1.76 **2.27**	1.74 **2.23**	1.73 **2.21**
25	4.24 **7.77**	3.38 **5.57**	2.99 **4.68**	2.76 **4.18**	2.60 **3.86**	2.49 **3.63**	2.41 **3.46**	2.34 **3.32**	2.28 **3.21**	2.24 **3.13**	2.20 **3.05**	2.16 **2.99**	2.11 **2.89**	2.06 **2.81**	2.00 **2.70**	1.96 **2.62**	1.92 **2.54**	1.87 **2.45**	1.84 **2.40**	1.80 **2.32**	1.77 **2.29**	1.74 **2.23**	1.72 **2.19**	1.71 **2.17**
26	4.22 **7.72**	3.37 **5.53**	2.89 **4.64**	2.74 **4.14**	2.59 **3.82**	2.47 **3.59**	2.39 **3.42**	2.32 **3.29**	2.27 **3.17**	2.22 **3.09**	2.18 **3.02**	2.15 **2.96**	2.10 **2.86**	2.05 **2.77**	1.99 **2.66**	1.95 **2.58**	1.90 **2.50**	1.85 **2.41**	1.82 **2.36**	1.78 **2.28**	1.76 **2.25**	1.72 **2.19**	1.70 **2.15**	1.69 **2.13**
27	4.21 **7.68**	3.35 **5.49**	2.96 **4.60**	2.73 **4.11**	2.57 **3.79**	2.46 **3.56**	2.37 **3.39**	2.30 **3.26**	2.25 **3.14**	2.20 **3.06**	2.16 **2.98**	2.13 **2.93**	2.08 **2.83**	2.03 **2.74**	1.97 **2.63**	1.93 **2.55**	1.88 **2.47**	1.84 **2.38**	1.80 **2.33**	1.76 **2.25**	1.74 **2.21**	1.71 **2.16**	1.68 **2.12**	1.67 **2.10**
28	4.20 **7.64**	3.34 **5.45**	2.95 **4.57**	2.71 **4.07**	2.56 **3.76**	2.44 **3.53**	2.36 **3.36**	2.29 **3.23**	3.24 **3.11**	2.19 **3.03**	2.15 **2.95**	2.12 **2.90**	2.06 **2.80**	2.02 **2.71**	1.96 **2.60**	1.91 **2.52**	1.87 **2.44**	1.81 **2.35**	1.78 **2.30**	1.75 **2.22**	1.72 **2.18**	1.69 **2.13**	1.67 **2.09**	1.65 **2.06**
29	4.18 **7.60**	3.33 **5.52**	2.93 **4.54**	2.70 **4.04**	2.54 **3.73**	2.43 **3.50**	2.35 **3.32**	2.28 **3.20**	2.22 **3.08**	2.18 **3.00**	2.14 **2.92**	2.10 **2.87**	2.05 **2.77**	2.00 **2.68**	1.94 **2.57**	1.90 **2.49**	1.85 **2.41**	1.80 **2.32**	1.77 **2.27**	1.73 **2.19**	1.71 **2.15**	1.68 **2.10**	1.65 **2.06**	1.64 **2.03**

Table B-7 Percentage points of the *F* distribution (Continued)

Degrees of freedom for lesser mean square																								
30	4.17 7.56	3.32 5.39	2.92 4.51	2.69 4.02	2.53 3.70	2.42 3.47	2.34 3.30	2.27 3.17	2.21 3.06	2.16 2.98	2.12 2.90	2.09 2.84	2.04 2.74	1.99 2.66	1.93 2.55	1.89 2.47	1.84 2.38	1.79 2.29	1.76 2.24	1.72 2.16	1.69 2.13	1.66 2.07	1.64 2.03	1.62 2.01
32	4.15 7.50	3.30 5.34	2.90 4.46	2.67 3.97	2.51 3.66	2.40 3.42	2.32 3.25	2.25 3.12	2.19 3.01	2.14 2.94	2.10 2.86	2.07 2.80	2.02 2.70	1.97 2.62	1.91 2.51	1.86 2.42	1.82 2.34	1.76 2.25	1.74 2.20	1.69 2.12	1.67 2.08	1.64 2.02	1.61 1.98	1.59 1.96
34	4.13 7.44	3.28 5.29	2.88 4.42	2.65 3.93	2.49 3.61	2.38 3.38	2.30 3.21	2.23 3.08	2.17 2.97	2.12 2.89	2.08 2.82	2.05 2.76	2.00 2.66	1.95 2.58	1.89 2.47	1.84 2.38	1.80 2.30	1.74 2.21	1.71 2.15	1.67 2.08	1.64 2.04	1.61 1.98	1.59 1.94	1.57 1.91
36	4.11 7.39	3.26 5.25	2.86 4.38	2.63 3.89	2.48 3.58	2.36 3.35	2.28 3.18	2.21 3.04	2.15 2.94	2.10 2.86	2.06 2.78	2.03 2.72	1.89 2.62	1.93 2.54	1.87 2.43	1.82 2.35	1.78 2.26	1.72 2.17	1.69 2.12	1.65 2.04	1.62 2.00	1.59 1.94	1.56 1.90	1.55 1.87
38	4.10 7.35	3.25 5.21	2.85 4.34	2.62 3.86	2.46 3.54	2.35 3.32	2.26 3.15	2.19 3.02	2.14 2.91	2.09 2.82	2.05 2.75	2.02 2.69	1.96 2.59	1.92 2.51	1.85 2.40	1.80 2.32	1.76 2.22	1.71 2.14	1.67 2.08	1.63 2.00	1.60 1.97	1.57 1.90	1.54 1.86	1.53 1.84
40	4.08 7.31	3.23 5.18	2.84 4.31	2.61 3.83	2.45 3.51	2.34 3.29	2.25 3.12	2.18 2.99	2.12 2.88	2.07 2.80	2.04 2.73	2.00 2.66	1.95 2.56	1.90 2.49	1.84 2.37	1.79 2.29	1.74 2.20	1.69 2.11	1.66 2.05	1.61 1.97	1.59 1.94	1.55 1.88	1.53 1.84	1.51 1.81
42	4.07 7.27	3.22 5.15	2.83 4.29	2.59 3.80	2.44 3.49	2.32 3.26	2.24 3.10	2.17 2.96	2.11 2.86	2.06 2.77	2.02 2.70	1.90 2.64	1.94 2.54	1.89 2.46	1.82 2.35	1.78 2.26	1.73 2.17	1.68 2.08	1.64 2.02	1.60 1.94	1.57 1.91	1.54 1.85	1.51 1.80	1.49 1.78
44	4.06 7.24	3.21 5.12	2.82 4.26	2.58 3.78	2.43 3.46	2.31 3.24	2.23 3.07	2.16 2.94	2.10 2.84	2.05 2.75	2.01 2.68	1.98 2.62	1.92 2.52	1.88 2.44	1.81 2.32	1.76 2.24	1.72 2.15	1.66 2.06	1.63 2.09	1.58 1.92	1.56 1.88	1.52 1.82	1.50 1.78	1.48 1.75
46	4.05 7.21	3.20 5.10	2.81 4.24	2.57 3.76	2.42 3.44	2.30 3.22	2.22 3.05	2.14 2.92	2.09 2.82	2.04 2.73	2.00 2.66	1.97 2.60	1.91 2.50	1.87 2.42	1.80 2.30	1.75 2.22	1.71 2.13	1.65 2.04	1.62 1.98	1.57 1.90	1.54 1.86	1.51 1.80	1.48 1.76	1.46 1.72
48	4.04 7.19	3.19 5.08	2.80 4.22	2.56 3.74	2.41 3.42	2.30 3.20	2.21 3.04	2.14 2.90	2.08 2.80	2.03 2.71	1.99 2.64	1.96 2.58	1.90 2.48	1.86 2.40	1.79 2.28	1.74 2.20	1.70 2.11	1.64 2.02	1.61 1.96	1.56 1.88	1.53 1.84	1.50 1.78	1.47 1.73	1.45 1.70
50	4.03 7.17	3.18 5.06	2.79 4.20	2.56 3.72	2.40 3.41	2.29 3.18	2.20 3.02	2.13 2.88	2.07 2.78	2.02 2.70	1.98 2.62	1.95 2.56	1.90 2.46	1.85 2.39	1.78 2.26	1.74 2.18	1.69 2.10	1.63 2.00	1.60 1.94	1.55 1.86	1.52 1.82	1.48 1.76	1.46 1.71	1.44 1.68
55	4.02 7.12	3.17 5.01	2.78 4.16	2.54 3.68	2.38 3.37	2.27 3.15	2.18 2.98	2.11 2.85	2.05 2.75	2.00 2.66	1.97 2.59	1.93 2.53	1.88 2.43	1.83 2.35	1.76 2.23	1.72 2.15	1.67 2.06	1.61 1.96	1.58 1.90	1.52 1.82	1.50 1.78	1.46 1.71	1.43 1.66	1.41 1.64
60	4.00 7.08	3.15 4.98	2.76 4.13	2.52 3.65	2.37 3.34	2.25 3.12	2.17 2.95	2.10 2.82	2.04 2.72	1.99 2.63	1.95 2.56	1.92 2.50	1.86 2.40	1.81 2.32	1.75 2.20	1.70 2.12	1.65 2.03	1.59 1.93	1.56 1.87	1.50 1.79	1.48 1.74	1.44 1.68	1.41 1.63	1.39 1.60
65	3.99 7.04	3.14 4.95	2.75 4.10	2.51 3.62	2.36 3.31	2.24 3.09	2.15 2.93	2.08 2.79	2.02 2.70	1.98 2.61	1.94 2.54	1.90 2.47	1.85 2.37	1.80 2.30	1.73 2.18	1.68 2.09	1.63 2.00	1.57 1.90	1.54 1.84	1.49 1.76	1.46 1.71	1.42 1.64	1.39 1.60	1.37 1.56
70	3.98 7.01	3.13 4.92	2.74 4.08	2.50 3.60	2.35 3.29	2.32 3.07	2.14 2.91	2.07 2.77	2.01 2.67	1.97 2.59	1.93 2.51	1.89 2.45	1.84 2.35	1.79 2.28	1.72 2.15	1.67 2.07	1.62 1.98	1.56 1.88	1.53 1.82	1.47 1.74	1.45 1.69	1.40 1.62	1.37 1.56	1.35 1.53
80	3.96 6.96	3.11 4.88	2.72 4.04	2.48 3.56	2.33 3.25	2.21 3.04	2.12 2.87	2.05 2.74	1.99 2.64	1.95 2.55	1.91 2.48	1.88 2.41	1.82 2.32	1.77 2.24	1.70 2.11	1.65 2.03	1.60 1.94	1.54 1.84	1.51 1.78	1.45 1.70	1.42 1.65	1.38 1.57	1.35 1.52	1.32 1.49

Table B-8 Binomial distribution

$$P_b(\bar{r} \geq r|n,p)$$

N = 1

R	P	.01	.02	.03	.04	.05	.06	.07	.08	.09	.10		
0		.9900	.9800	.9700	.9600	.9500	.9400	.9300	.9200	.9100	.9000		1
1		.0100	.0200	.0300	.0400	.0500	.0600	.0700	.0800	.0900	.1000		0
		.99	.98	.97	.96	.95	.94	.93	.92	.91	.90	P	R

R	P	.11	.12	.13	.14	.15	.16	.17	.18	.19	.20		
0		.8900	.8800	.8700	.8600	.8500	.8400	.8300	.8200	.8100	.8000		1
1		.1100	.1200	.1300	.1400	.1500	.1600	.1700	.1800	.1900	.2000		0
		.89	.88	.87	.86	.85	.84	.83	.82	.81	.80	P	R

R	P	.21	.22	.23	.24	.25	.26	.27	.28	.29	.30		
0		.7900	.7800	.7700	.7600	.7500	.7400	.7300	.7200	.7100	.7000		1
1		.2100	.2200	.2300	.2400	.2500	.2600	.2700	.2800	.2900	.3000		0
		.79	.78	.77	.76	.75	.74	.73	.72	.71	.70	P	R

R	P	.31	.32	.33	.34	.35	.36	.37	.38	.39	.40		
0		.6900	.6800	.6700	.6600	.6500	.6400	.6300	.6200	.6100	.6000		1
1		.3100	.3200	.3300	.3400	.3500	.3600	.3700	.3800	.3900	.4000		0
		.69	.68	.67	.66	.65	.64	.63	.62	.61	.60	P	R

R	P	.41	.42	.43	.44	.45	.46	.47	.48	.49	.50		
0		.5900	.5800	.5700	.5600	.5500	.5400	.5300	.5200	.5100	.5000		1
1		.4100	.4200	.4300	.4400	.4500	.4600	.4700	.4800	.4900	.5000		0
		.59	.58	.57	.56	.55	.54	.53	.52	.51	.50	P	R

R	P	.01	.02	.03	.04	.05	.06	.07	.08	.09	.10		
0		.9801	.9604	.9409	.9216	.9025	.8836	.8649	.8464	.8281	.8100		2
1		.0198	.0392	.0582	.0768	.0950	.1128	.1302	.1472	.1638	.1800		1
2		.0001	.0004	.0009	.0016	.0025	.0036	.0049	.0064	.0081	.0100		0
		.99	.98	.97	.96	.95	.94	.93	.92	.91	.90	P	R

R	P	.11	.12	.13	.14	.15	.16	.17	.18	.19	.20		
0		.7921	.7744	.7569	.7396	.7225	.7056	.6889	.6724	.6561	.6400		2
1		.1958	.2112	.2262	.2408	.2550	.2688	.2822	.2952	.3078	.3200		1
2		.0121	.0144	.0169	.0196	.0225	.0256	.0289	.0324	.0361	.0400		0
		.89	.88	.87	.86	.85	.84	.83	.82	.81	.80	P	R

R	P	.21	.22	.23	.24	.25	.26	.27	.28	.29	.30		
0		.6241	.6084	.5929	.5776	.5625	.5476	.5329	.5184	.5041	.4900		2
1		.3318	.3432	.3542	.3648	.3750	.3848	.3942	.4032	.4118	.4200		1
2		.0441	.0484	.0529	.0576	.0625	.0676	.0729	.0784	.0841	.0900		0
		.79	.78	.77	.76	.75	.74	.73	.72	.71	.70	P	R

R	P	.31	.32	.33	.34	.35	.36	.37	.38	.39	.40		
0		.4761	.4624	.4489	.4356	.4225	.4096	.3969	.3844	.3721	.3600		2
1		.4278	.4352	.4422	.4488	.4550	.4608	.4662	.4712	.4758	.4800		1
2		.0961	.1024	.1089	.1156	.1225	.1296	.1369	.1444	.1521	.1600		0
		.69	.68	.67	.66	.65	.64	.63	.62	.61	.60	P	R

R	P	.41	.42	.43	.44	.45	.46	.47	.48	.49	.50		
0		.3481	.3364	.3249	.3136	.3025	.2916	.2809	.2704	.2601	.2500		2
1		.4838	.4872	.4902	.4928	.4950	.4968	.4982	.4992	.4998	.5000		1
2		.1681	.1764	.1849	.1936	.2025	.2116	.2209	.2304	.2401	.2500		0
		.59	.58	.57	.56		.54						

From Robert A. Parsons, *Statistical Analysis: A Decision Making Approach*, Harper & Row, Publishers, Incorporated, New York, forthcoming 1974.

Table B-8 Binomial distribution (Continued)

R	P	.01	.02	.03	.04	.05	.06	.07	.08	.09	.10		
0		.9703	.9412	.9127	.8847	.8574	.8306	.8044	.7787	.7536	.7290		3
1		.0294	.0576	.0847	.1106	.1354	.1590	.1816	.2031	.2236	.2430		2
2		.0003	.0012	.0026	.0046	.0071	.0102	.0137	.0177	.0221	.0270		1
3		.0000	.0000	.0000	.0001	.0001	.0002	.0003	.0005	.0007	.0010		0
		.99	.98	.97	.96	.95	.94	.93	.92	.91	.90	P	R

R	P	.11	.12	.13	.14	.15	.16	.17	.18	.19	.20		
0		.7050	.6815	.6585	.6361	.6141	.5927	.5718	.5514	.5314	.5120		3
1		.2614	.2788	.2952	.3106	.3251	.3387	.3513	.3631	.3740	.3840		2
2		.0323	.0380	.0441	.0506	.0574	.0645	.0720	.0797	.0877	.0960		1
3		.0013	.0017	.0022	.0027	.0034	.0041	.0049	.0058	.0069	.0080		0
		.89	.88	.87	.86	.85	.84	.83	.82	.81	.80	P	R

R	P	.21	.22	.23	.24	.25	.26	.27	.28	.29	.30		
0		.4930	.4746	.4565	.4390	.4219	.4052	.3890	.3732	.3579	.3430		3
1		.3932	.4015	.4091	.4159	.4219	.4271	.4316	.4355	.4386	.4410		2
2		.1045	.1133	.1222	.1313	.1406	.1501	.1597	.1693	.1791	.1890		1
3		.0093	.0106	.0122	.0138	.0156	.0176	.0197	.0220	.0244	.0270		0
		.79	.78	.77	.76	.75	.74	.73	.72	.71	.70	P	R

R	P	.31	.32	.33	.34	.35	.36	.37	.38	.39	.40		
0		.3285	.3144	.3008	.2875	.2746	.2621	.2500	.2383	.2270	.2160		3
1		.4428	.4439	.4444	.4443	.4436	.4424	.4406	.4382	.4354	.4320		2
2		.1989	.2089	.2189	.2289	.2389	.2488	.2587	.2686	.2783	.2880		1
3		.0298	.0328	.0359	.0393	.0429	.0467	.0507	.0549	.0593	.0640		0
		.69	.68	.67	.66	.65	.64	.63	.62	.61	.60	P	R

R	P	.41	.42	.43	.44	.45	.46	.47	.48	.49	.50		
0		.2054	.1951	.1852	.1756	.1664	.1575	.1489	.1406	.1327	.1250		3
1		.4282	.4239	.4191	.4140	.4084	.4024	.3961	.3894	.3823	.3750		2
2		.2975	.3069	.3162	.3252	.3341	.3428	.3512	.3594	.3674	.3750		1
3		.0689	.0741	.0795	.0852	.0911	.0973	.1038	.1106	.1176	.1250		0
		.59	.58	.57	.56	.55	.54	.53	.52	.51	.50	P	R

R	P	.01	.02	.03	.04	.05	.06	.07	.08	.09	.10		
0		.9606	.9224	.8853	.8493	.8145	.7807	.7481	.7164	.6857	.6561		4
1		.0388	.0753	.1095	.1416	.1715	.1993	.2252	.2492	.2713	.2916		3
2		.0006	.0023	.0051	.0088	.0135	.0191	.0254	.0325	.0402	.0486		2
3		.0000	.0000	.0001	.0002	.0005	.0008	.0013	.0019	.0027	.0036		1
4		.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0001		0
		.99	.98	.97	.96	.95	.94	.93	.92	.91	.90	P	R

R	P	.11	.12	.13	.14	.15	.16	.17	.18	.19	.20		
0		.6274	.5997	.5729	.5470	.5220	.4979	.4746	.4521	.4305	.4096		4
1		.3102	.3271	.3424	.3562	.3685	.3793	.3888	.3970	.4039	.4096		3
2		.0575	.0669	.0767	.0870	.0975	.1084	.1195	.1307	.1421	.1536		2
3		.0047	.0061	.0076	.0094	.0115	.0138	.0163	.0191	.0222	.0256		1
4		.0001	.0002	.0003	.0004	.0005	.0007	.0008	.0010	.0013	.0016		0
		.89	.88	.87	.86	.85	.84	.83	.82	.81	.80	P	R

R	P	.21	.22	.23	.24	.25	.26	.27	.28	.29	.30		
0		.3895	.3702	.3515	.3336	.3164	.2999	.2840	.2687	.2541	.2401		4
1		.4142	.4176	.4200	.4214	.4219	.4214	.4201	.4180	.4152	.4116		3
2		.1651	.1767	.1882	.1996	.2109	.2221	.2331	.2439	.2544	.2646		2
3		.0293	.0332	.0375	.0420	.0469	.0520	.0575	.0632	.0693	.0756		1
4		.0019	.0023	.0028	.0033	.0039	.0046	.0053	.0061	.0071	.0081		0
		.79	.78	.77	.76	.75	.74	.73	.72	.71	.70	P	R

Table B-8 Binomial distribution (Continued)

R	P	.31	.32	.33	.34	.35	.36	.37	.38	.39	.40		
0		.2267	.2138	.2015	.1897	.1785	.1678	.1575	.1478	.1385	.1296		4
1		.4074	.4025	.3970	.3910	.3845	.3775	.3701	.3623	.3541	.3456		3
2		.2745	.2841	.2933	.3021	.3105	.3185	.3260	.3330	.3396	.3456		2
3		.0822	.0891	.0963	.1038	.1115	.1194	.1276	.1361	.1447	.1536		1
4		.0092	.0105	.0119	.0134	.0150	.0168	.0187	.0209	.0231	.0256		0
		.69	.68	.67	.66	.65	.64	.63	.62	.61	.60	P	R

R	P	.41	.42	.43	.44	.45	.46	.47	.48	.49	.50		
0		.1212	.1132	.1056	.0983	.0915	.0850	.0789	.0731	.0677	.0625		4
1		.3368	.3278	.3185	.3091	.2995	.2897	.2799	.2700	.2600	.2500		3
2		.3511	.3560	.3604	.3643	.3675	.3702	.3723	.3738	.3747	.3750		2
3		.1627	.1719	.1813	.1908	.2005	.2102	.2201	.2300	.2400	.2500		1
4		.0283	.0311	.0342	.0375	.0410	.0448	.0488	.0531	.0576	.0625		0
		.59	.58	.57	.56	.55	.54	.53	.52	.51	.50	P	R

R	P	.01	.02	.03	.04	.05	.06	.07	.08	.09	.10		
0		.9510	.9039	.8587	.8154	.7738	.7339	.6957	.6591	.6240	.5905		5
1		.0480	.0922	.1328	.1699	.2036	.2342	.2618	.2866	.3086	.3281		4
2		.0010	.0038	.0082	.0142	.0214	.0299	.0394	.0498	.0610	.0729		3
3		.0000	.0001	.0003	.0006	.0011	.0019	.0030	.0043	.0060	.0081		2
4		.0000	.0000	.0000	.0000	.0000	.0001	.0001	.0002	.0003	.0005		1
		.99	.98	.97	.96	.95	.94	.93	.92	.91	.90	P	R

R	P	.11	.12	.13	.14	.15	.16	.17	.18	.19	.20		
0		.5584	.5277	.4984	.4704	.4437	.4182	.3939	.3707	.3487	.3277		5
1		.3451	.3598	.3724	.3829	.3915	.3983	.4034	.4069	.4089	.4096		4
2		.0853	.0981	.1113	.1247	.1382	.1517	.1652	.1786	.1919	.2048		3
3		.0105	.0134	.0166	.0203	.0244	.0289	.0338	.0392	.0450	.0512		2
4		.0007	.0009	.0012	.0017	.0022	.0028	.0035	.0043	.0053	.0064		1
5		.0000	.0000	.0000	.0001	.0001	.0001	.0001	.0002	.0002	.0003		0
		.89	.88	.87	.86	.85	.84	.83	.82	.81	.80	P	R

R	P	.21	.22	.23	.24	.25	.26	.27	.28	.29	.30		
0		.3077	.2887	.2707	.2536	.2373	.2219	.2073	.1935	.1804	.1681		5
1		.4090	.4072	.4043	.4003	.3955	.3898	.3834	.3762	.3685	.3602		4
2		.2174	.2297	.2415	.2529	.2637	.2739	.2836	.2926	.3010	.3087		3
3		.0578	.0648	.0721	.0798	.0879	.0962	.1049	.1138	.1229	.1323		2
4		.0077	.0091	.0108	.0126	.0146	.0169	.0194	.0221	.0251	.0284		1
5		.0004	.0005	.0006	.0008	.0010	.0012	.0014	.0017	.0021	.0024		0
		.79	.78	.77	.76	.75	.74	.73	.72	.71	.70	P	R

R	P	.31	.32	.33	.34	.35	.36	.37	.38	.39	.40		
0		.1564	.1454	.1350	.1252	.1160	.1074	.0992	.0916	.0845	.0778		5
1		.3513	.3421	.3325	.3226	.3124	.3020	.2914	.2808	.2700	.2592		4
2		.3157	.3220	.3275	.3323	.3364	.3397	.3423	.3441	.3452	.3456		3
3		.1418	.1515	.1613	.1712	.1811	.1911	.2010	.2109	.2207	.2304		2
4		.0319	.0357	.0397	.0441	.0488	.0537	.0590	.0646	.0706	.0768		1
5		.0029	.0034	.0039	.0045	.0053	.0060	.0069	.0079	.0090	.0102		0
		.69	.68	.67	.66	.65	.64	.63	.62	.61	.60	P	R

R	P	.41	.42	.43	.44	.45	.46	.47	.48	.49	.50		
0		.0715	.0656	.0602	.0551	.0503	.0459	.0418	.0380	.0345	.0313		5
1		.2484	.2376	.2270	.2164	.2059	.1956	.1854	.1755	.1657	.1563		4
2		.3452	.3442	.3424	.3400	.3369	.3332	.3289	.3240	.3185	.3125		3
3		.2399	.2492	.2583	.2671	.2757	.2838	.2916	.2990	.3060	.3125		2
4		.0834	.0902	.0974	.1049	.1128	.1209	.1293	.1380	.1470	.1563		1
5		.0116	.0131	.0147	.0165	.0185	.0206	.0229	.0255	.0282	.0313		0
		.59	.58	.57	.56	.55	.54	.53	.52	.51	.50	P	R

Table B-8 Binomial distribution (Continued)

R	P	.01	.02	.03	.04	.05	.06	.07	.08	.09	.10		
0		.9415	.8858	.8330	.7828	.7351	.6899	.6470	.6064	.5679	.5314		6
1		.0571	.1085	.1546	.1957	.2321	.2642	.2922	.3164	.3370	.3543		5
2		.0014	.0055	.0120	.0204	.0305	.0422	.0550	.0688	.0833	.0984		4
3		.0000	.0002	.0005	.0011	.0021	.0036	.0055	.0080	.0110	.0146		3
4		.0000	.0000	.0000	.0000	.0001	.0002	.0003	.0005	.0008	.0012		2
5		.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001		1
		.99	.98	.97	.96	.95	.94	.93	.92	.91	.90	P	R

R	P	.11	.12	.13	.14	.15	.16	.17	.18	.19	.20		
0		.4970	.4644	.4336	.4046	.3771	.3513	.3269	.3040	.2824	.2621		6
1		.3685	.3800	.3888	.3952	.3993	.4015	.4018	.4004	.3975	.3932		5
2		.1139	.1295	.1452	.1608	.1762	.1912	.2057	.2197	.2331	.2458		4
3		.0188	.0236	.0289	.0349	.0415	.0486	.0562	.0643	.0729	.0819		3
4		.0017	.0024	.0032	.0043	.0055	.0069	.0086	.0106	.0128	.0154		2
5		.0001	.0001	.0002	.0003	.0004	.0005	.0007	.0009	.0012	.0015		1
6		.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001		0
		.89	.88	.87	.86	.85	.84	.83	.82	.81	.80	P	R

R	P	.21	.22	.23	.24	.25	.26	.27	.28	.29	.30		
0		.2431	.2252	.2084	.1927	.1780	.1642	.1513	.1393	.1281	.1176		6
1		.3877	.3811	.3735	.3651	.3560	.3462	.3358	.3251	.3139	.3025		5
2		.2577	.2687	.2789	.2882	.2966	.3041	.3105	.3160	.3206	.3241		4
3		.0913	.1011	.1111	.1214	.1318	.1424	.1531	.1639	.1746	.1852		3
4		.0182	.0214	.0249	.0287	.0330	.0375	.0425	.0478	.0535	.0595		2
5		.0019	.0024	.0030	.0036	.0044	.0053	.0063	.0074	.0087	.0102		1
6		.0001	.0001	.0001	.0002	.0002	.0003	.0004	.0005	.0006	.0007		0
		.79	.78	.77	.76	.75	.74	.73	.72	.71	.70	P	R

R	P	.31	.32	.33	.34	.35	.36	.37	.38	.39	.40		
0		.1079	.0989	.0905	.0827	.0754	.0687	.0625	.0568	.0515	.0467		6
1		.2909	.2792	.2673	.2555	.2437	.2319	.2203	.2089	.1976	.1866		5
2		.3267	.3284	.3292	.3290	.3280	.3261	.3235	.3201	.3159	.3110		4
3		.1957	.2061	.2162	.2260	.2355	.2446	.2533	.2616	.2693	.2765		3
4		.0660	.0727	.0799	.0873	.0951	.1032	.1116	.1202	.1291	.1382		2
5		.0119	.0137	.0157	.0180	.0205	.0232	.0262	.0295	.0330	.0369		1
6		.0009	.0011	.0013	.0015	.0018	.0022	.0026	.0030	.0035	.0041		0
		.69	.68	.67	.66	.65	.64	.63	.62	.61	.60	P	R

R	P	.41	.42	.43	.44	.45	.46	.47	.48	.49	.50		
0		.0422	.0381	.0343	.0308	.0277	.0248	.0222	.0198	.0176	.0156		6
1		.1759	.1654	.1552	.1454	.1359	.1267	.1179	.1095	.1014	.0938		5
2		.3055	.2994	.2928	.2856	.2780	.2699	.2615	.2527	.2436	.2344		4
3		.2831	.2891	.2945	.2992	.3032	.3065	.3091	.3110	.3121	.3125		3
4		.1475	.1570	.1666	.1763	.1861	.1958	.2056	.2153	.2249	.2344		2
5		.0410	.0455	.0503	.0554	.0609	.0667	.0729	.0795	.0864	.0938		1
6		.0048	.0055	.0063	.0073	.0083	.0095	.0108	.0122	.0138	.0156		0
		.59	.58	.57	.56	.55	.54	.53	.52	.51	.50	P	R

R	P	.01	.02	.03	.04	.05	.06	.07	.08	.09	.10		
0		.9321	.8681	.8080	.7514	.6983	.6485	.6017	.5578	.5168	.4783		7
1		.0659	.1240	.1749	.2192	.2573	.2897	.3170	.3396	.3578	.3720		6
2		.0020	.0076	.0162	.0274	.0406	.0555	.0716	.0886	.1061	.1240		5
3		.0000	.0003	.0008	.0019	.0036	.0059	.0090	.0128	.0175	.0230		4
4		.0000	.0000	.0000	.0001	.0002	.0004	.0007	.0011	.0017	.0026		3
5		.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0001	.0002		2
		.99	.98	.97	.96	.95	.94	.93	.92	.91	.90	P	R

Table B-8 Binomial distribution (Continued)

R	P	.11	.12	.13	.14	.15	.16	.17	.18	.19	.20		
0		.4423	.4087	.3773	.3479	.3205	.2951	.2714	.2493	.2288	.2097		7
1		.3827	.3901	.3946	.3965	.3960	.3935	.3891	.3830	.3756	.3670		6
2		.1419	.1596	.1769	.1936	.2097	.2248	.2391	.2523	.2643	.2753		5
3		.0292	.0363	.0441	.0525	.0617	.0714	.0816	.0923	.1033	.1147		4
4		.0036	.0049	.0066	.0086	.0109	.0136	.0167	.0203	.0242	.0287		3
5		.0003	.0004	.0006	.0008	.0012	.0016	.0021	.0027	.0034	.0043		2
6		.0000	.0000	.0000	.0000	.0001	.0001	.0001	.0002	.0003	.0004		1
		.89	.88	.87	.86	.85	.84	.83	.82	.81	.80	P	R

R	P	.21	.22	.23	.24	.25	.26	.27	.28	.29	.30		
0		.1920	.1757	.1605	.1465	.1335	.1215	.1105	.1003	.0910	.0824		7
1		.3573	.3468	.3356	.3237	.3115	.2989	.2860	.2731	.2600	.2471		6
2		.2850	.2935	.3007	.3067	.3115	.3150	.3174	.3186	.3186	.3177		5
3		.1263	.1379	.1497	.1614	.1730	.1845	.1956	.2065	.2169	.2269		4
4		.0336	.0389	.0447	.0510	.0577	.0648	.0724	.0803	.0886	.0972		3
5		.0054	.0066	.0080	.0097	.0115	.0137	.0161	.0187	.0217	.0250		2
6		.0005	.0006	.0008	.0010	.0013	.0016	.0020	.0024	.0030	.0036		1
7		.0000	.0000	.0000	.0000	.0001	.0001	.0001	.0001	.0002	.0002		0
		.79	.78	.77	.76	.75	.74	.73	.72	.71	.70	P	R

R	P	.31	.32	.33	.34	.35	.36	.37	.38	.39	.40		
0		.0745	.0672	.0606	.0546	.0490	.0440	.0394	.0352	.0314	.0280		7
1		.2342	.2215	.2090	.1967	.1848	.1732	.1619	.1511	.1407	.1306		6
2		.3156	.3127	.3088	.3040	.2985	.2922	.2853	.2778	.2698	.2613		5
3		.2363	.2452	.2535	.2610	.2679	.2740	.2793	.2838	.2875	.2903		4
4		.1062	.1154	.1248	.1345	.1442	.1541	.1640	.1739	.1838	.1935		3
5		.0286	.0326	.0369	.0416	.0466	.0520	.0578	.0640	.0705	.0774		2
6		.0043	.0051	.0061	.0071	.0084	.0098	.0113	.0131	.0150	.0172		1
7		.0003	.0003	.0004	.0005	.0006	.0008	.0009	.0011	.0014	.0016		0
		.69	.68	.67	.66	.65	.64	.63	.62	.61	.60	P	R

R	P	.41	.42	.43	.44	.45	.46	.47	.48	.49	.50		
0		.0249	.0221	.0195	.0173	.0152	.0134	.0117	.0103	.0090	.0078		7
1		.1211	.1119	.1032	.0950	.0872	.0798	.0729	.0664	.0604	.0547		6
2		.2524	.2431	.2336	.2239	.2140	.2040	.1940	.1840	.1740	.1641		5
3		.2923	.2934	.2937	.2932	.2918	.2897	.2867	.2830	.2786	.2734		4
4		.2031	.2125	.2216	.2304	.2388	.2468	.2543	.2612	.2676	.2734		3
5		.0847	.0923	.1003	.1086	.1172	.1261	.1353	.1447	.1543	.1641		2
6		.0196	.0223	.0252	.0284	.0320	.0358	.0400	.0445	.0494	.0547		1
7		.0019	.0023	.0027	.0032	.0037	.0044	.0051	.0059	.0068	.0078		0
		.59	.58	.57	.56	.55	.54	.53	.52	.51	.50	P	R

R	P	.01	.02	.03	.04	.05	.06	.07	.08	.09	.10		
0		.9227	.8508	.7837	.7214	.6634	.6096	.5596	.5132	.4703	.4305		8
1		.0746	.1389	.1939	.2405	.2793	.3113	.3370	.3570	.3721	.3826		7
2		.0026	.0099	.0210	.0351	.0515	.0695	.0888	.1087	.1288	.1488		6
3		.0001	.0004	.0013	.0029	.0054	.0089	.0134	.0189	.0255	.0331		5
4		.0000	.0000	.0001	.0002	.0004	.0007	.0013	.0021	.0031	.0046		4
5		.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0001	.0002	.0004		3
		.99	.98	.97	.96	.95	.94	.93	.92	.91	.90	P	R

R	P	.11	.12	.13	.14	.15	.16	.17	.18	.19	.20		
0		.3937	.3596	.3282	.2992	.2725	.2479	.2252	.2044	.1853	.1678		8
1		.3892	.3923	.3923	.3897	.3847	.3777	.3691	.3590	.3477	.3355		7
2		.1684	.1872	.2052	.2220	.2376	.2518	.2646	.2758	.2855	.2936		6
3		.0416	.0511	.0613	.0723	.0839	.0959	.1084	.1211	.1339	.1468		5
4		.0064	.0087	.0115	.0147	.0185	.0228	.0277	.0332	.0393	.0459		4
5		.0006	.0009	.0014	.0019	.0026	.0035	.0045	.0058	.0074	.0092		3
6		.0000	.0001	.0001	.0002	.0002	.0003	.0005	.0006	.0009	.0011		2
7		.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0001		1
		.89	.88	.87	.86	.85	.84	.83	.82	.81	.80	P	R

Table B-8 Binomial distribution (Continued)

R	P	.21	.22	.23	.24	.25	.26	.27	.28	.29	.30		
0		.1517	.1370	.1236	.1113	.1001	.0899	.0806	.0722	.0646	.0576		8
1		.3226	.3092	.2953	.2812	.2670	.2527	.2386	.2247	.2110	.1977		7
2		.3002	.3052	.3087	.3108	.3115	.3108	.3089	.3058	.3017	.2965		6
3		.1596	.1722	.1844	.1963	.2076	.2184	.2285	.2379	.2464	.2541		5
4		.0530	.0607	.0689	.0775	.0865	.0959	.1056	.1156	.1258	.1361		4
5		.0113	.0137	.0165	.0196	.0231	.0270	.0313	.0360	.0411	.0467		3
6		.0015	.0019	.0025	.0031	.0038	.0047	.0058	.0070	.0084	.0100		2
7		.0001	.0002	.0002	.0003	.0004	.0005	.0006	.0008	.0010	.0012		1
8		.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0001		0
		.79	.78	.77	.76	.75	.74	.73	.72	.71	.70	P	R

R	P	.31	.32	.33	.34	.35	.36	.37	.38	.39	.40		
0		.0514	.0457	.0406	.0360	.0319	.0281	.0248	.0218	.0192	.0168		8
1		.1847	.1721	.1600	.1484	.1373	.1267	.1166	.1071	.0981	.0896		7
2		.2904	.2835	.2758	.2675	.2587	.2494	.2397	.2297	.2194	.2090		6
3		.2609	.2668	.2717	.2756	.2786	.2805	.2815	.2815	.2806	.2787		5
4		.1465	.1569	.1673	.1775	.1875	.1973	.2067	.2157	.2242	.2322		4
5		.0527	.0591	.0659	.0732	.0808	.0888	.0971	.1058	.1147	.1239		3
6		.0118	.0139	.0162	.0188	.0217	.0250	.0285	.0324	.0367	.0413		2
7		.0015	.0019	.0023	.0028	.0033	.0040	.0048	.0057	.0067	.0079		1
8		.0001	.0001	.0001	.0002	.0002	.0003	.0004	.0004	.0005	.0007		0
		.69	.68	.67	.66	.65	.64	.63	.62	.61	.60	P	R

R	P	.41	.42	.43	.44	.45	.46	.47	.48	.49	.50		
0		.0147	.0128	.0111	.0097	.0084	.0072	.0062	.0053	.0046	.0039		8
1		.0816	.0742	.0672	.0608	.0548	.0493	.0442	.0395	.0352	.0313		7
2		.1985	.1880	.1776	.1672	.1569	.1469	.1371	.1275	.1183	.1094		6
3		.2759	.2723	.2679	.2627	.2568	.2503	.2431	.2355	.2273	.2188		5
4		.2397	.2465	.2526	.2580	.2627	.2665	.2695	.2717	.2730	.2734		4
5		.1332	.1428	.1525	.1622	.1719	.1816	.1912	.2006	.2098	.2188		3
6		.0463	.0517	.0575	.0637	.0703	.0774	.0848	.0926	.1008	.1094		2
7		.0092	.0107	.0124	.0143	.0164	.0188	.0215	.0244	.0277	.0313		1
8		.0008	.0010	.0012	.0014	.0017	.0020	.0024	.0028	.0033	.0039		0
		.59	.58	.57	.56	.55	.54	.53	.52	.51	.50	P	R

R	P	.01	.02	.03	.04	.05	.06	.07	.08	.09	.10		
0		.9135	.8337	.7602	.6925	.6302	.5730	.5204	.4722	.4279	.3874		9
1		.0830	.1531	.2116	.2597	.2985	.3292	.3525	.3695	.3809	.3874		8
2		.0034	.0125	.0262	.0433	.0629	.0840	.1061	.1285	.1507	.1722		7
3		.0001	.0006	.0019	.0042	.0077	.0125	.0186	.0261	.0348	.0446		6
4		.0000	.0000	.0001	.0003	.0006	.0012	.0021	.0034	.0052	.0074		5
5		.0000	.0000	.0000	.0000	.0000	.0001	.0002	.0003	.0005	.0008		4
6		.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001		3
		.99	.98	.97	.96	.95	.94	.93	.92	.91	.90	P	R

R	P	.11	.12	.13	.14	.15	.16	.17	.18	.19	.20		
0		.3504	.3165	.2855	.2573	.2316	.2082	.1869	.1676	.1501	.1342		9
1		.3897	.3884	.3840	.3770	.3679	.3569	.3446	.3312	.3169	.3020		8
2		.1927	.2119	.2295	.2455	.2597	.2720	.2823	.2908	.2973	.3020		7
3		.0556	.0674	.0800	.0933	.1069	.1209	.1349	.1489	.1627	.1762		6
4		.0103	.0138	.0179	.0228	.0283	.0345	.0415	.0490	.0573	.0661		5
5		.0013	.0019	.0027	.0037	.0050	.0066	.0085	.0108	.0134	.0165		4
6		.0001	.0002	.0003	.0004	.0006	.0008	.0012	.0016	.0021	.0028		3
7		.0000	.0000	.0000	.0000	.0000	.0001	.0001	.0001	.0002	.0003		2
		.89	.88	.87	.86	.85	.84	.83	.82	.81	.80	P	R

R	P	.21	.22	.23	.24	.25	.26	.27	.28	.29	.30		
0		.1199	.1069	.0952	.0846	.0751	.0665	.0589	.0520	.0458	.0404		9
1		.2867	.2713	.2558	.2404	.2253	.2104	.1950	.1820	.1685	.1556		8
2		.3049	.3061	.3056	.3037	.3003	.2957	.2899	.2831	.2754	.2668		7
3		.1891	.2014	.2130	.2238	.2336	.2424	.2502	.2569	.2624	.2668		6

Table B-8 Binomial distribution (Continued)

4		.0754	.0852	.0954	.1060	.1168	.1278	.1388	.1499	.1608	.1715		5
5		.0200	.0240	.0285	.0335	.0389	.0449	.0513	.0583	.0657	.0735		4
6		.0036	.0045	.0057	.0070	.0087	.0105	.0127	.0151	.0179	.0210		3
7		.0004	.0005	.0007	.0010	.0012	.0016	.0020	.0025	.0031	.0039		2
8		.0000	.0000	.0001	.0001	.0001	.0001	.0002	.0002	.0003	.0004		1
		.79	.78	.77	.76	.75	.74	.73	.72	.71	.70	P	R

R	P	.31	.32	.33	.34	.35	.36	.37	.38	.39	.40		
0		.0355	.0311	.0272	.0238	.0207	.0180	.0156	.0135	.0117	.0101		9
1		.1433	.1317	.1206	.1102	.1004	.0912	.0826	.0747	.0673	.0605		8
2		.2576	.2478	.2376	.2270	.2162	.2052	.1941	.1831	.1721	.1612		7
3		.2701	.2721	.2731	.2729	.2716	.2693	.2660	.2618	.2567	.2508		6
4		.1820	.1921	.2017	.2109	.2194	.2272	.2344	.2407	.2462	.2508		5
5		.0818	.0904	.0994	.1086	.1181	.1278	.1376	.1475	.1574	.1672		4
6		.0245	.0284	.0326	.0373	.0424	.0479	.0539	.0603	.0671	.0743		3
7		.0047	.0057	.0069	.0082	.0098	.0116	.0136	.0156	.0184	.0212		2
8		.0005	.0007	.0008	.0011	.0013	.0016	.0020	.0024	.0029	.0035		1
9		.0000	.0000	.0000	.0001	.0001	.0001	.0001	.0002	.0002	.0003		0
		.69	.68	.67	.66	.65	.64	.63	.62	.61	.60	P	R

R	P	.41	.42	.43	.44	.45	.46	.47	.48	.49	.50		
0		.0087	.0074	.0064	.0054	.0046	.0039	.0033	.0028	.0023	.0020		9
1		.0542	.0484	.0431	.0383	.0339	.0299	.0263	.0231	.0202	.0176		8
2		.1506	.1402	.1301	.1204	.1110	.1020	.0934	.0853	.0776	.0703		7
3		.2442	.2369	.2291	.2207	.2119	.2027	.1933	.1837	.1739	.1641		6
4		.2545	.2573	.2592	.2601	.2600	.2590	.2571	.2543	.2506	.2461		5
5		.1769	.1863	.1955	.2044	.2128	.2207	.2280	.2347	.2408	.2461		4
6		.0819	.0900	.0983	.1070	.1160	.1253	.1348	.1445	.1542	.1641		3
7		.0244	.0279	.0318	.0360	.0407	.0458	.0512	.0571	.0635	.0703		2
8		.0042	.0051	.0060	.0071	.0083	.0097	.0114	.0132	.0153	.0176		1
9		.0003	.0004	.0005	.0006	.0008	.0009	.0011	.0014	.0016	.0020		0
		.59	.58	.57	.56	.55	.54	.53	.52	.51	.50	P	R

10

R	P	.01	.02	.03	.04	.05	.06	.07	.08	.09	.10		
0		.9044	.8171	.7374	.6648	.5987	.5386	.4840	.4344	.3894	.3487		10
1		.0914	.1667	.2281	.2770	.3151	.3438	.3643	.3777	.3851	.3874		9
2		.0042	.0153	.0317	.0519	.0746	.0988	.1234	.1478	.1714	.1937		8
3		.0001	.0008	.0026	.0058	.0105	.0168	.0248	.0343	.0452	.0574		7
4		.0000	.0000	.0001	.0004	.0010	.0019	.0033	.0052	.0078	.0112		6
5		.0000	.0000	.0000	.0000	.0001	.0001	.0003	.0005	.0009	.0015		5
6		.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0001		4
		.99	.98	.97	.96	.95	.94	.93	.92	.91	.90	P	R

R	P	.11	.12	.13	.14	.15	.16	.17	.18	.19	.20		
0		.3118	.2785	.2484	.2213	.1969	.1749	.1552	.1374	.1216	.1074		10
1		.3854	.3798	.3712	.3603	.3474	.3331	.3178	.3017	.2852	.2684		9
2		.2143	.2330	.2496	.2639	.2759	.2856	.2929	.2980	.3010	.3020		8
3		.0706	.0847	.0995	.1146	.1298	.1450	.1600	.1745	.1883	.2013		7
4		.0153	.0202	.0260	.0326	.0401	.0483	.0573	.0670	.0773	.0881		6
5		.0023	.0033	.0047	.0064	.0085	.0111	.0141	.0177	.0218	.0264		5
6		.0002	.0004	.0006	.0009	.0012	.0018	.0024	.0032	.0043	.0055		4
7		.0000	.0000	.0000	.0001	.0001	.0002	.0003	.0004	.0006	.0008		3
8		.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0001		2
		.89	.88	.87	.86	.85	.84	.83	.82	.81	.80	P	R

R	P	.21	.22	.23	.24	.25	.26	.27	.28	.29	.30		
0		.0947	.0834	.0733	.0643	.0563	.0492	.0430	.0374	.0326	.0282		10
1		.2517	.2351	.2188	.2030	.1877	.1730	.1590	.1456	.1330	.1211		9
2		.3011	.2984	.2942	.2885	.2816	.2735	.2646	.2548	.2444	.2335		8
3		.2134	.2244	.2343	.2429	.2503	.2563	.2609	.2642	.2662	.2668		7
4		.0993	.1108	.1225	.1343	.1460	.1576	.1689	.1798	.1903	.2001		6
5		.0317	.0375	.0439	.0509	.0584	.0664	.0750	.0839	.0933	.1029		5

Table B-8 Binomial distribution (Continued)

6		.0070	.0088	.0109	.0134	.0162	.0195	.0231	.0272	.0317	.0368		4
7		.0011	.0014	.0019	.0024	.0031	.0039	.0049	.0060	.0074	.0090		3
8		.0001	.0002	.0002	.0003	.0004	.0005	.0007	.0009	.0011	.0014		2
9		.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0001	.0001	.0001		1
		.79	.78	.77	.76	.75	.74	.73	.72	.71	.70	P	R

R	P	.31	.32	.33	.34	.35	.36	.37	.38	.39	.40		
0		.0245	.0211	.0182	.0157	.0135	.0115	.0098	.0084	.0071	.0060		10
1		.1099	.0995	.0898	.0808	.0725	.0649	.0578	.0514	.0456	.0403		9
2		.2222	.2107	.1990	.1873	.1757	.1642	.1529	.1419	.1312	.1209		8
3		.2662	.2644	.2614	.2573	.2522	.2462	.2394	.2319	.2237	.2150		7
4		.2093	.2177	.2253	.2320	.2377	.2424	.2461	.2487	.2503	.2508		6
5		.1128	.1229	.1332	.1434	.1536	.1636	.1734	.1829	.1920	.2007		5
6		.0422	.0482	.0547	.0616	.0689	.0767	.0849	.0934	.1023	.1115		4
7		.0108	.0130	.0154	.0181	.0212	.0247	.0285	.0327	.0374	.0425		3
8		.0018	.0023	.0028	.0035	.0043	.0052	.0063	.0075	.0090	.0106		2
9		.0002	.0002	.0003	.0004	.0005	.0007	.0008	.0010	.0013	.0016		1
10		.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0001	.0001		0
		.69	.68	.67	.66	.65	.64	.63	.62	.61	.60	P	R

R	P	.41	.42	.43	.44	.45	.46	.47	.48	.49	.50		
0		.0051	.0043	.0036	.0030	.0025	.0021	.0017	.0014	.0012	.0010		10
1		.0355	.0312	.0273	.0238	.0207	.0180	.0155	.0133	.0114	.0098		9
2		.1111	.1017	.0927	.0843	.0763	.0688	.0619	.0554	.0495	.0439		8
3		.2058	.1963	.1865	.1765	.1665	.1564	.1464	.1364	.1267	.1172		7
4		.2503	.2488	.2462	.2427	.2384	.2331	.2271	.2204	.2130	.2051		6
5		.2087	.2162	.2229	.2289	.2340	.2383	.2417	.2441	.2456	.2461		5
6		.1209	.1304	.1401	.1499	.1596	.1692	.1786	.1878	.1966	.2051		4
7		.0480	.0540	.0604	.0673	.0746	.0824	.0905	.0991	.1080	.1172		3
8		.0125	.0147	.0171	.0198	.0229	.0263	.0301	.0343	.0389	.0439		2
9		.0019	.0024	.0029	.0035	.0042	.0050	.0059	.0070	.0083	.0098		1
10		.0001	.0002	.0002	.0003	.0003	.0004	.0005	.0006	.0008	.0010		0
		.59	.58	.57	.56	.55	.54	.53	.52	.51	.50	P	R

N = 11

R	P	.01	.02	.03	.04	.05	.06	.07	.08	.09	.10		
0		.8953	.8007	.7153	.6382	.5688	.5063	.4501	.3996	.3544	.3138		11
1		.0995	.1798	.2434	.2925	.3293	.3555	.3727	.3823	.3855	.3835		10
2		.0050	.0183	.0376	.0609	.0867	.1135	.1403	.1662	.1906	.2131		9
3		.0002	.0011	.0035	.0076	.0137	.0217	.0317	.0434	.0566	.0710		8
4		.0000	.0000	.0002	.0006	.0014	.0028	.0048	.0075	.0112	.0158		7
5		.0000	.0000	.0000	.0000	.0001	.0002	.0005	.0009	.0015	.0025		6
6		.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0002	.0003		5
		.99	.98	.97	.96	.95	.94	.93	.92	.91	.90	P	R

R	P	.11	.12	.13	.14	.15	.16	.17	.18	.19	.20		
0		.2775	.2451	.2161	.1903	.1673	.1469	.1288	.1127	.0985	.0859		11
1		.3773	.3676	.3552	.3408	.3248	.3078	.2902	.2721	.2541	.2362		10
2		.2332	.2507	.2654	.2774	.2866	.2932	.2971	.2987	.2980	.2953		9
3		.0865	.1025	.1190	.1355	.1517	.1675	.1826	.1967	.2097	.2215		8
4		.0214	.0280	.0356	.0441	.0536	.0638	.0748	.0864	.0984	.1107		7
5		.0037	.0053	.0074	.0101	.0132	.0170	.0214	.0265	.0323	.0388		6
6		.0005	.0007	.0011	.0016	.0023	.0032	.0044	.0058	.0076	.0097		5
7		.0000	.0001	.0001	.0002	.0003	.0004	.0006	.0009	.0013	.0017		4
8		.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0001	.0001	.0002		3
		.89	.88	.87	.86	.85	.84	.83	.82	.81	.80	P	R

R	P	.21	.22	.23	.24	.25	.26	.27	.28	.29	.30		
0		.0748	.0650	.0564	.0489	.0422	.0364	.0314	.0270	.0231	.0198		11
1		.2187	.2017	.1854	.1697	.1549	.1408	.1276	.1153	.1038	.0932		10
2		.2907	.2845	.2768	.2680	.2581	.2474	.2360	.2242	.2121	.1998		9
3		.2318	.2407	.2481	.2539	.2581	.2608	.2619	.2616	.2599	.2568		8
4		.1232	.1358	.1482	.1603	.1721	.1832	.1937	.2035	.2123	.2201		7

Table B-8 Binomial distribution (Continued)

5		.0459	.0536	.0620	.0709	.0803	.0901	.1003	.1108	.1214	.1321		6
6		.0122	.0151	.0185	.0224	.0268	.0317	.0371	.0431	.0496	.0566		5
7		.0023	.0030	.0039	.0050	.0064	.0079	.0098	.0120	.0145	.0173		4
8		.0003	.0004	.0006	.0008	.0011	.0014	.0018	.0023	.0030	.0037		3
9		.0000	.0000	.0001	.0001	.0001	.0002	.0002	.0003	.0004	.0005		2
		.79	.78	.77	.76	.75	.74	.73	.72	.71	.70	P	R

R	P	.31	.32	.33	.34	.35	.36	.37	.38	.39	.40		
0		.0169	.0144	.0122	.0104	.0088	.0074	.0062	.0052	.0044	.0036		11
1		.0834	.0744	.0662	.0587	.0518	.0457	.0401	.0351	.0306	.0266		10
2		.1874	.1751	.1630	.1511	.1395	.1284	.1177	.1075	.0978	.0887		9
3		.2526	.2472	.2408	.2335	.2254	.2167	.2074	.1977	.1876	.1774		8
4		.2269	.2326	.2372	.2406	.2428	.2438	.2436	.2423	.2399	.2365		7
5		.1427	.1533	.1636	.1735	.1830	.1920	.2003	.2079	.2148	.2207		6
6		.0641	.0721	.0806	.0894	.0985	.1080	.1176	.1274	.1373	.1471		5
7		.0206	.0242	.0283	.0329	.0379	.0434	.0494	.0558	.0627	.0701		4
8		.0046	.0057	.0070	.0085	.0102	.0122	.0145	.0171	.0200	.0234		3
9		.0007	.0009	.0011	.0015	.0018	.0023	.0028	.0035	.0043	.0052		2
10		.0001	.0001	.0001	.0001	.0002	.0003	.0003	.0004	.0005	.0007		1
		.69	.68	.67	.66	.65	.64	.63	.62	.61	.60	P	R

R	P	.41	.42	.43	.44	.45	.46	.47	.48	.49	.50		
0		.0030	.0025	.0021	.0017	.0014	.0011	.0009	.0008	.0006	.0005		11
1		.0231	.0199	.0171	.0147	.0125	.0107	.0090	.0076	.0064	.0054		10
2		.0801	.0721	.0646	.0577	.0513	.0454	.0401	.0352	.0308	.0269		9
3		.1670	.1566	.1462	.1359	.1259	.1161	.1067	.0976	.0888	.0806		8
4		.2321	.2267	.2206	.2136	.2060	.1978	.1892	.1801	.1707	.1611		7
5		.2258	.2299	.2329	.2350	.2360	.2359	.2348	.2327	.2296	.2256		6
6		.1569	.1664	.1757	.1846	.1931	.2010	.2083	.2148	.2206	.2256		5
7		.0779	.0861	.0947	.1036	.1128	.1223	.1319	.1416	.1514	.1611		4
8		.0271	.0312	.0357	.0407	.0462	.0521	.0585	.0654	.0727	.0806		3
9		.0063	.0075	.0090	.0107	.0126	.0148	.0173	.0201	.0233	.0269		2
10		.0009	.0011	.0014	.0017	.0021	.0025	.0031	.0037	.0045	.0054		1
11		.0001	.0001	.0001	.0001	.0002	.0002	.0002	.0003	.0004	.0005		0
		.59	.58	.57	.56	.55	.54	.53	.52	.51	.50	P	R

N = 12

R	P	.01	.02	.03	.04	.05	.06	.07	.08	.09	.10		
0		.8864	.7847	.6938	.6127	.5404	.4759	.4186	.3677	.3225	.2824		12
1		.1074	.1922	.2575	.3064	.3413	.3645	.3781	.3837	.3827	.3766		11
2		.0060	.0216	.0438	.0702	.0988	.1280	.1565	.1835	.2082	.2301		10
3		.0002	.0015	.0045	.0098	.0173	.0272	.0393	.0532	.0686	.0852		9
4		.0000	.0001	.0003	.0009	.0021	.0039	.0067	.0104	.0153	.0213		8
5		.0000	.0000	.0000	.0001	.0002	.0004	.0008	.0014	.0024	.0038		7
6		.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0001	.0003	.0005		6
		.99	.98	.97	.96	.95	.94	.93	.92	.91	.90	P	R

R	P	.11	.12	.13	.14	.15	.16	.17	.18	.19	.20		
0		.2470	.2157	.1880	.1637	.1422	.1234	.1069	.0924	.0798	.0687		12
1		.3663	.3529	.3372	.3197	.3012	.2821	.2627	.2434	.2245	.2062		11
2		.2490	.2647	.2771	.2863	.2924	.2955	.2960	.2939	.2897	.2835		10
3		.1026	.1203	.1380	.1553	.1720	.1876	.2021	.2151	.2265	.2362		9
4		.0285	.0369	.0464	.0569	.0683	.0804	.0931	.1062	.1195	.1329		8
5		.0056	.0081	.0111	.0148	.0193	.0245	.0305	.0373	.0449	.0532		7
6		.0008	.0013	.0019	.0028	.0040	.0054	.0073	.0096	.0123	.0155		6
7		.0001	.0001	.0002	.0004	.0006	.0009	.0013	.0018	.0025	.0033		5
8		.0000	.0000	.0000	.0000	.0001	.0001	.0002	.0002	.0004	.0005		4
9		.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001		3
		.89	.88	.87	.86	.85	.84	.83	.82	.81	.80	P	R

R	P	.21	.22	.23	.24	.25	.26	.27	.28	.29	.30		
0		.0591	.0507	.0434	.0371	.0317	.0270	.0229	.0194	.0164	.0138		12
1		.1885	.1717	.1557	.1407	.1267	.1137	.1016	.0906	.0804	.0712		11

Table B-8 Binomial distribution (Continued)

2		.2756	.2663	.2558	.2444	.2323	.2197	.2068	.1937	.1807	.1678		10
3		.2442	.2503	.2547	.2573	.2581	.2573	.2549	.2511	.2460	.2397		9
4		.1460	.1589	.1712	.1828	.1936	.2034	.2122	.2197	.2261	.2311		8
5		.0621	.0717	.0818	.0924	.1032	.1143	.1255	.1367	.1477	.1585		7
6		.0193	.0236	.0285	.0340	.0401	.0469	.0542	.0620	.0704	.0792		6
7		.0044	.0057	.0073	.0092	.0115	.0141	.0172	.0207	.0246	.0291		5
8		.0007	.0010	.0014	.0018	.0024	.0031	.0040	.0050	.0063	.0078		4
9		.0001	.0001	.0002	.0003	.0004	.0005	.0007	.0009	.0011	.0015		3
10		.0000	.0000	.0000	.0000	.0000	.0001	.0001	.0001	.0001	.0002		2
		.79	.78	.77	.76	.75	.74	.73	.72	.71	.70	P	R

R	P	.31	.32	.33	.34	.35	.36	.37	.38	.39	.40		
0		.0116	.0098	.0082	.0068	.0057	.0047	.0039	.0032	.0027	.0022		12
1		.0628	.0552	.0484	.0422	.0368	.0319	.0276	.0237	.0204	.0174		11
2		.1552	.1429	.1310	.1197	.1088	.0986	.0890	.0800	.0716	.0639		10
3		.2324	.2241	.2151	.2055	.1954	.1849	.1742	.1634	.1526	.1419		9
4		.2349	.2373	.2384	.2382	.2367	.2340	.2302	.2254	.2195	.2128		8
5		.1688	.1787	.1879	.1963	.2039	.2106	.2163	.2210	.2246	.2270		7
6		.0885	.0981	.1079	.1180	.1281	.1382	.1482	.1580	.1675	.1766		6
7		.0341	.0396	.0456	.0521	.0591	.0666	.0746	.0830	.0918	.1009		5
8		.0096	.0116	.0140	.0168	.0199	.0234	.0274	.0318	.0367	.0420		4
9		.0019	.0024	.0031	.0038	.0048	.0059	.0071	.0087	.0104	.0125		3
10		.0003	.0003	.0005	.0006	.0008	.0010	.0013	.0016	.0020	.0025		2
11		.0000	.0000	.0000	.0001	.0001	.0001	.0001	.0002	.0002	.0003		1
		.69	.68	.67	.66	.65	.64	.63	.62	.61	.60	P	R

R	P	.41	.42	.43	.44	.45	.46	.47	.48	.49	.50		
0		.0018	.0014	.0012	.0010	.0008	.0006	.0005	.0004	.0003	.0002		12
1		.0148	.0126	.0106	.0090	.0075	.0063	.0052	.0043	.0036	.0029		11
2		.0567	.0502	.0442	.0388	.0339	.0294	.0255	.0220	.0189	.0161		10
3		.1314	.1211	.1111	.1015	.0923	.0836	.0754	.0676	.0604	.0537		9
4		.2054	.1973	.1886	.1794	.1700	.1602	.1504	.1405	.1306	.1208		8
5		.2284	.2285	.2276	.2256	.2225	.2184	.2134	.2075	.2008	.1934		7
6		.1851	.1931	.2003	.2068	.2124	.2171	.2208	.2234	.2250	.2256		6
7		.1103	.1198	.1295	.1393	.1489	.1585	.1678	.1768	.1853	.1934		5
8		.0479	.0542	.0611	.0684	.0762	.0844	.0930	.1020	.1113	.1208		4
9		.0148	.0175	.0205	.0239	.0277	.0319	.0367	.0418	.0475	.0537		3
10		.0031	.0038	.0046	.0056	.0068	.0082	.0098	.0116	.0137	.0161		2
11		.0004	.0005	.0006	.0008	.0010	.0013	.0016	.0019	.0024	.0029		1
12		.0000	.0000	.0000	.0001	.0001	.0001	.0001	.0001	.0002	.0002		0
		.59	.58	.57	.56	.55	.54	.53	.52	.51	.50	P	R

N = 13

R	P	.01	.02	.03	.04	.05	.06	.07	.08	.09	.10		
0		.8775	.7690	.6730	.5882	.5133	.4474	.3893	.3383	.2935	.2542		13
1		.1152	.2040	.2706	.3186	.3512	.3712	.3809	.3824	.3773	.3672		12
2		.0070	.0250	.0502	.0797	.1109	.1422	.1720	.1995	.2239	.2448		11
3		.0003	.0019	.0057	.0122	.0214	.0333	.0475	.0636	.0812	.0997		10
4		.0000	.0001	.0004	.0013	.0028	.0053	.0089	.0138	.0201	.0277		9
5		.0000	.0000	.0000	.0001	.0003	.0006	.0012	.0022	.0036	.0055		8
6		.0000	.0000	.0000	.0000	.0000	.0001	.0001	.0003	.0005	.0008		7
7		.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001		6
		.99	.98	.97	.96	.95	.94	.93	.92	.91	.90	P	R

R	P	.11	.12	.13	.14	.15	.16	.17	.18	.19	.20		
0		.2198	.1898	.1636	.1408	.1209	.1037	.0887	.0758	.0646	.0550		13
1		.3532	.3364	.3178	.2979	.2774	.2567	.2362	.2163	.1970	.1787		12
2		.2619	.2753	.2849	.2910	.2937	.2934	.2903	.2848	.2773	.2680		11
3		.1187	.1376	.1561	.1737	.1900	.2049	.2180	.2293	.2385	.2457		10
4		.0367	.0469	.0583	.0707	.0838	.0976	.1116	.1258	.1399	.1535		9
5		.0082	.0115	.0157	.0207	.0266	.0335	.0412	.0497	.0591	.0691		8
6		.0013	.0021	.0031	.0045	.0063	.0085	.0112	.0145	.0185	.0230		7
7		.0002	.0003	.0005	.0007	.0011	.0016	.0023	.0032	.0043	.0058		6

Table B-8 Binomial distribution (Continued)

8		.0000	.0000	.0001	.0001	.0001	.0002	.0004	.0005	.0008	.0011		5
9		.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0001	.0002		4
		.89	.88	.87	.86	.85	.84	.83	.82	.81	.80	P	R

R	P	.21	.22	.23	.24	.25	.26	.27	.28	.29	.30		
0		.0467	.0396	.0334	.0282	.0238	.0200	.0167	.0140	.0117	.0097		13
1		.1613	.1450	.1299	.1159	.1029	.0911	.0804	.0706	.0619	.0540		12
2		.2573	.2455	.2328	.2195	.2059	.1921	.1784	.1648	.1516	.1388		11
3		.2508	.2539	.2550	.2542	.2517	.2475	.2419	.2351	.2271	.2181		10
4		.1667	.1790	.1904	.2007	.2097	.2174	.2237	.2285	.2319	.2337		9
5		.0797	.0909	.1024	.1141	.1258	.1375	.1489	.1600	.1705	.1803		8
6		.0283	.0342	.0408	.0480	.0559	.0644	.0734	.0829	.0928	.1030		7
7		.0075	.0096	.0122	.0152	.0186	.0226	.0272	.0323	.0379	.0442		6
8		.0015	.0020	.0027	.0036	.0047	.0060	.0075	.0094	.0116	.0142		5
9		.0002	.0003	.0005	.0006	.0009	.0012	.0015	.0020	.0026	.0034		4
10		.0000	.0000	.0001	.0001	.0001	.0002	.0002	.0003	.0004	.0006		3
11		.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001		2
		.79	.78	.77	.76	.75	.74	.73	.72	.71	.70	P	R

R	P	.31	.32	.33	.34	.35	.36	.37	.38	.39	.40		
0		.0080	.0066	.0055	.0045	.0037	.0030	.0025	.0020	.0016	.0013		13
1		.0469	.0407	.0351	.0302	.0259	.0221	.0188	.0159	.0135	.0113		12
2		.1265	.1148	.1037	.0933	.0836	.0746	.0663	.0586	.0516	.0453		11
3		.2084	.1981	.1874	.1763	.1651	.1538	.1427	.1317	.1210	.1107		10
4		.2341	.2331	.2307	.2270	.2222	.2163	.2095	.2018	.1934	.1845		9
5		.1893	.1974	.2045	.2105	.2154	.2190	.2215	.2227	.2226	.2214		8
6		.1134	.1239	.1343	.1446	.1546	.1643	.1734	.1820	.1898	.1968		7
7		.0510	.0583	.0662	.0745	.0833	.0924	.1019	.1115	.1213	.1312		6
8		.0172	.0206	.0244	.0288	.0336	.0390	.0449	.0513	.0582	.0656		5
9		.0043	.0054	.0067	.0082	.0101	.0122	.0146	.0175	.0207	.0243		4
10		.0008	.0010	.0013	.0017	.0022	.0027	.0034	.0043	.0053	.0065		3
11		.0001	.0001	.0002	.0002	.0003	.0004	.0006	.0007	.0009	.0012		2
12		.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0001	.0001	.0001		1
		.69	.68	.67	.66	.65	.64	.63	.62	.61	.60	P	R

R	P	.41	.42	.43	.44	.45	.46	.47	.48	.49	.50		
0		.0010	.0008	.0007	.0005	.0004	.0003	.0003	.0002	.0002	.0001		13
1		.0095	.0079	.0066	.0054	.0045	.0037	.0030	.0024	.0020	.0016		12
2		.0395	.0344	.0298	.0256	.0220	.0188	.0160	.0135	.0114	.0095		11
3		.1007	.0913	.0823	.0739	.0660	.0587	.0519	.0457	.0401	.0349		10
4		.1750	.1653	.1553	.1451	.1350	.1250	.1151	.1055	.0962	.0873		9
5		.2189	.2154	.2108	.2053	.1989	.1917	.1838	.1753	.1664	.1571		8
6		.2029	.2080	.2121	.2151	.2169	.2177	.2173	.2158	.2131	.2095		7
7		.1410	.1506	.1600	.1690	.1775	.1854	.1927	.1992	.2048	.2095		6
8		.0735	.0818	.0905	.0996	.1089	.1185	.1282	.1379	.1476	.1571		5
9		.0284	.0329	.0379	.0435	.0495	.0561	.0631	.0707	.0788	.0873		4
10		.0079	.0095	.0114	.0137	.0162	.0191	.0224	.0261	.0303	.0349		3
11		.0015	.0019	.0024	.0029	.0036	.0044	.0054	.0066	.0079	.0095		2
12		.0002	.0002	.0003	.0004	.0005	.0006	.0008	.0010	.0013	.0016		1
13		.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0001	.0001	.0001		0
		.59	.58	.57	.56	.55	.54	.53	.52	.51	.50	P	R

N = 14

R	P	.01	.02	.03	.04	.05	.06	.07	.08	.09	.10		
0		.8687	.7536	.6528	.5647	.4877	.4205	.3620	.3112	.2670	.2288		14
1		.1229	.2153	.2827	.3294	.3593	.3758	.3815	.3788	.3698	.3559		13
2		.0081	.0286	.0568	.0892	.1229	.1559	.1867	.2141	.2377	.2570		12
3		.0003	.0023	.0070	.0149	.0259	.0398	.0562	.0745	.0940	.1142		11
4		.0000	.0001	.0006	.0017	.0037	.0070	.0116	.0178	.0256	.0349		10
5		.0000	.0000	.0000	.0001	.0004	.0009	.0018	.0031	.0051	.0078		9
6		.0000	.0000	.0000	.0000	.0000	.0001	.0002	.0004	.0008	.0013		8
7		.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0002		7
		.99	.98	.97	.96	.95	.94	.93	.92	.91	.90	P	R

Table B-8 Binomial distribution (Continued)

R	P	.11	.12	.13	.14	.15	.16	.17	.18	.19	.20		
0		.1956	.1670	.1423	.1211	.1028	.0871	.0736	.0621	.0523	.0440		14
1		.3385	.3188	.2977	.2759	.2539	.2322	.2112	.1910	.1719	.1539		13
2		.2720	.2826	.2892	.2919	.2912	.2875	.2811	.2725	.2620	.2501		12
3		.1345	.1542	.1728	.1901	.2056	.2190	.2303	.2393	.2459	.2501		11
4		.0457	.0578	.0710	.0851	.0998	.1147	.1297	.1444	.1586	.1720		10
5		.0113	.0158	.0212	.0277	.0352	.0437	.0531	.0634	.0744	.0860		9
6		.0021	.0032	.0048	.0068	.0093	.0125	.0163	.0209	.0262	.0322		8
7		.0003	.0005	.0008	.0013	.0019	.0027	.0038	.0052	.0070	.0092		7
8		.0000	.0001	.0001	.0002	.0003	.0005	.0007	.0010	.0014	.0020		6
9		.0000	.0000	.0000	.0000	.0000	.0001	.0001	.0001	.0002	.0003		5
		.89	.88	.87	.86	.85	.84	.83	.82	.81	.80	P	R

R	P	.21	.22	.23	.24	.25	.26	.27	.28	.29	.30		
0		.0369	.0309	.0258	.0214	.0178	.0148	.0122	.0101	.0083	.0068		14
1		.1372	.1218	.1077	.0948	.0832	.0726	.0632	.0548	.0473	.0407		13
2		.2371	.2234	.2091	.1946	.1802	.1659	.1519	.1385	.1256	.1134		12
3		.2521	.2520	.2499	.2459	.2402	.2331	.2248	.2154	.2052	.1943		11
4		.1843	.1955	.2052	.2135	.2202	.2252	.2286	.2304	.2305	.2290		10
5		.0980	.1103	.1226	.1348	.1468	.1583	.1691	.1792	.1883	.1963		9
6		.0391	.0466	.0549	.0639	.0734	.0834	.0938	.1045	.1153	.1262		8
7		.0119	.0150	.0188	.0231	.0280	.0335	.0397	.0464	.0538	.0618		7
8		.0028	.0037	.0049	.0064	.0082	.0103	.0128	.0158	.0192	.0232		6
9		.0005	.0007	.0010	.0013	.0018	.0024	.0032	.0041	.0052	.0066		5
10		.0001	.0001	.0001	.0002	.0003	.0004	.0006	.0008	.0011	.0014		4
11		.0000	.0000	.0000	.0000	.0000	.0001	.0001	.0001	.0002	.0002		3
		.79	.78	.77	.76	.75	.74	.73	.72	.71	.70	P	R

R	P	.31	.32	.33	.34	.35	.36	.37	.38	.39	.40		
0		.0055	.0045	.0037	.0030	.0024	.0019	.0016	.0012	.0010	.0008		14
1		.0349	.0298	.0253	.0215	.0181	.0152	.0128	.0106	.0088	.0073		13
2		.1018	.0911	.0811	.0719	.0634	.0557	.0487	.0424	.0367	.0317		12
3		.1830	.1715	.1598	.1481	.1366	.1253	.1144	.1039	.0940	.0845		11
4		.2261	.2219	.2164	.2098	.2022	.1938	.1848	.1752	.1652	.1549		10
5		.2032	.2088	.2132	.2161	.2178	.2181	.2170	.2147	.2112	.2066		9
6		.1369	.1474	.1575	.1670	.1759	.1840	.1912	.1974	.2026	.2066		8
7		.0703	.0793	.0886	.0983	.1082	.1183	.1283	.1383	.1480	.1574		7
8		.0276	.0326	.0382	.0443	.0510	.0582	.0659	.0742	.0828	.0918		6
9		.0083	.0102	.0125	.0152	.0183	.0218	.0258	.0303	.0353	.0408		5
10		.0019	.0024	.0031	.0039	.0049	.0061	.0076	.0093	.0113	.0136		4
11		.0003	.0004	.0006	.0007	.0010	.0013	.0016	.0021	.0026	.0033		3
12		.0000	.0000	.0001	.0001	.0001	.0002	.0002	.0003	.0004	.0005		2
13		.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001		1
		.69	.68	.67	.66	.65	.64	.63	.62	.61	.60	P	R

R	P	.41	.42	.43	.44	.45	.46	.47	.48	.49	.50		
0		.0006	.0005	.0004	.0003	.0002	.0002	.0001	.0001	.0001	.0001		14
1		.0060	.0049	.0040	.0033	.0027	.0021	.0017	.0014	.0011	.0009		13
2		.0272	.0233	.0198	.0168	.0141	.0118	.0099	.0082	.0068	.0056		12
3		.0757	.0674	.0597	.0527	.0462	.0403	.0350	.0303	.0260	.0222		11
4		.1446	.1342	.1239	.1138	.1040	.0945	.0854	.0768	.0687	.0611		10
5		.2009	.1943	.1869	.1788	.1701	.1610	.1515	.1418	.1320	.1222		9
6		.2094	.2111	.2115	.2108	.2088	.2057	.2015	.1963	.1902	.1833		8
7		.1663	.1747	.1824	.1892	.1952	.2003	.2043	.2071	.2089	.2095		7
8		.1011	.1107	.1204	.1301	.1398	.1493	.1585	.1673	.1756	.1833		6
9		.0469	.0534	.0605	.0682	.0762	.0848	.0937	.1030	.1125	.1222		5
10		.0163	.0193	.0228	.0268	.0312	.0361	.0415	.0475	.0540	.0611		4
11		.0041	.0051	.0063	.0076	.0093	.0112	.0134	.0160	.0189	.0222		3
12		.0007	.0009	.0012	.0015	.0019	.0024	.0030	.0037	.0045	.0056		2
13		.0001	.0001	.0001	.0002	.0002	.0003	.0004	.0005	.0007	.0009		1
14		.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001		0
		.59	.58	.57	.56	.55	.54	.53	.52	.51	.50	P	R

Table B-8 Binomial distribution (Continued)

N = 15

R	P	.01	.02	.03	.04	.05	.06	.07	.08	.09	.10		
0		.8601	.7386	.6333	.5421	.4633	.3953	.3367	.2863	.2430	.2059		15
1		.1303	.2261	.2938	.3388	.3658	.3785	.3801	.3734	.3605	.3432		14
2		.0092	.0323	.0636	.0988	.1348	.1691	.2003	.2273	.2496	.2669		13
3		.0004	.0029	.0085	.0178	.0307	.0468	.0653	.0857	.1070	.1285		12
4		.0000	.0002	.0008	.0022	.0049	.0090	.0148	.0223	.0317	.0428		11
5		.0000	.0000	.0001	.0002	.0006	.0013	.0024	.0043	.0069	.0105		10
6		.0000	.0000	.0000	.0000	.0000	.0001	.0003	.0006	.0011	.0019		9
7		.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0001	.0003		8
		.99	.98	.97	.96	.95	.94	.93	.92	.91	.90	P	R

R	P	.11	.12	.13	.14	.15	.16	.17	.18	.19	.20		
0		.1741	.1470	.1238	.1041	.0874	.0731	.0611	.0510	.0424	.0352		15
1		.3228	.3006	.2775	.2542	.2312	.2090	.1878	.1678	.1492	.1319		14
2		.2793	.2870	.2903	.2897	.2856	.2787	.2692	.2578	.2449	.2309		13
3		.1496	.1696	.1880	.2044	.2184	.2300	.2389	.2452	.2489	.2501		12
4		.0555	.0694	.0843	.0998	.1156	.1314	.1468	.1615	.1752	.1876		11
5		.0151	.0208	.0277	.0357	.0449	.0551	.0662	.0780	.0904	.1032		10
6		.0031	.0047	.0069	.0097	.0132	.0175	.0226	.0285	.0353	.0430		9
7		.0005	.0008	.0013	.0020	.0030	.0043	.0059	.0081	.0107	.0138		8
8		.0001	.0001	.0002	.0003	.0005	.0008	.0012	.0018	.0025	.0035		7
9		.0000	.0000	.0000	.0000	.0001	.0001	.0002	.0003	.0005	.0007		6
10		.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0001		5
		.89	.88	.87	.86	.85	.84	.83	.82	.81	.80	P	R

R	P	.21	.22	.23	.24	.25	.26	.27	.28	.29	.30		
0		.0291	.0241	.0198	.0163	.0134	.0109	.0089	.0072	.0059	.0047		15
1		.1162	.1018	.0889	.0772	.0668	.0576	.0494	.0423	.0360	.0305		14
2		.2162	.2010	.1858	.1707	.1559	.1416	.1280	.1150	.1029	.0916		13
3		.2490	.2457	.2405	.2336	.2252	.2156	.2051	.1939	.1821	.1700		12
4		.1986	.2079	.2155	.2213	.2252	.2273	.2276	.2262	.2231	.2186		11
5		.1161	.1290	.1416	.1537	.1651	.1757	.1852	.1935	.2005	.2061		10
6		.0514	.0606	.0705	.0809	.0917	.1029	.1142	.1254	.1365	.1472		9
7		.0176	.0220	.0271	.0329	.0393	.0465	.0543	.0627	.0717	.0811		8
8		.0047	.0062	.0081	.0104	.0131	.0163	.0201	.0244	.0293	.0348		7
9		.0010	.0014	.0019	.0025	.0034	.0045	.0058	.0074	.0093	.0116		6
10		.0002	.0002	.0003	.0005	.0007	.0009	.0013	.0017	.0023	.0030		5
11		.0000	.0000	.0000	.0001	.0001	.0002	.0002	.0003	.0004	.0006		4
12		.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0001		3
		.79	.78	.77	.76	.75	.74	.73	.72	.71	.70	P	R

R	P	.31	.32	.33	.34	.35	.36	.37	.38	.39	.40		
0		.0038	.0031	.0025	.0020	.0016	.0012	.0010	.0008	.0006	.0005		15
1		.0258	.0217	.0182	.0152	.0126	.0104	.0086	.0071	.0058	.0047		14
2		.0811	.0715	.0627	.0547	.0476	.0411	.0354	.0303	.0259	.0219		13
3		.1579	.1457	.1338	.1222	.1110	.1002	.0901	.0805	.0716	.0634		12
4		.2128	.2057	.1977	.1888	.1792	.1692	.1587	.1481	.1374	.1268		11
5		.2103	.2130	.2142	.2140	.2123	.2093	.2051	.1997	.1933	.1859		10
6		.1575	.1671	.1759	.1837	.1906	.1963	.2008	.2040	.2059	.2066		9
7		.0910	.1011	.1114	.1217	.1319	.1419	.1516	.1608	.1693	.1771		8
8		.0409	.0476	.0549	.0627	.0710	.0798	.0890	.0985	.1082	.1181		7
9		.0143	.0174	.0210	.0251	.0298	.0349	.0407	.0470	.0538	.0612		6
10		.0038	.0049	.0062	.0078	.0096	.0118	.0143	.0173	.0206	.0245		5
11		.0008	.0011	.0014	.0018	.0024	.0030	.0038	.0048	.0060	.0074		4
12		.0001	.0002	.0002	.0003	.0004	.0006	.0007	.0010	.0013	.0016		3
13		.0000	.0000	.0000	.0000	.0001	.0001	.0001	.0001	.0002	.0003		2
		.69	.68	.67	.66	.65	.64	.63	.62	.61	.60	P	R

Table B-8 Binomial distribution (Continued)

R	P	.41	.42	.43	.44	.45	.46	.47	.48	.49	.50		
0		.0004	.0003	.0002	.0002	.0001	.0001	.0001	.0001	.0000	.0000		15
1		.0038	.0031	.0025	.0020	.0016	.0012	.0010	.0008	.0006	.0005		14
2		.0185	.0156	.0130	.0108	.0090	.0074	.0060	.0049	.0040	.0032		13
3		.0558	.0489	.0426	.0369	.0318	.0272	.0232	.0197	.0166	.0139		12
4		.1163	.1061	.0963	.0869	.0780	.0696	.0617	.0545	.0478	.0417		11
5		.1778	.1691	.1598	.1502	.1404	.1304	.1204	.1106	.1010	.0916		10
6		.2060	.2041	.2010	.1967	.1914	.1851	.1780	.1702	.1617	.1527		9
7		.1840	.1900	.1949	.1987	.2013	.2028	.2030	.2020	.1997	.1964		8
8		.1279	.1376	.1470	.1561	.1647	.1727	.1800	.1864	.1919	.1964		7
9		.0691	.0775	.0863	.0954	.1048	.1144	.1241	.1338	.1434	.1527		6
10		.0288	.0337	.0390	.0450	.0515	.0585	.0661	.0741	.0827	.0916		5
11		.0091	.0111	.0134	.0161	.0191	.0226	.0266	.0311	.0361	.0417		4
12		.0021	.0027	.0034	.0042	.0052	.0064	.0079	.0096	.0116	.0139		3
13		.0003	.0004	.0006	.0008	.0010	.0013	.0016	.0020	.0026	.0032		2
14		.0000	.0000	.0001	.0001	.0001	.0002	.0002	.0003	.0004	.0005		1
		.59	.58	.57	.56	.55	.54	.53	.52	.51	.50	P	R

N = 16

R	P	.01	.02	.03	.04	.05	.06	.07	.08	.09	.10		
0		.8515	.7238	.6143	.5204	.4401	.3716	.3131	.2634	.2211	.1853		16
1		.1376	.2363	.3040	.3469	.3706	.3795	.3771	.3665	.3499	.3294		15
2		.0104	.0362	.0705	.1084	.1463	.1817	.2129	.2390	.2596	.2745		14
3		.0005	.0034	.0102	.0211	.0359	.0541	.0748	.0970	.1198	.1423		13
4		.0000	.0002	.0010	.0029	.0061	.0112	.0183	.0274	.0385	.0514		12
5		.0000	.0000	.0001	.0003	.0008	.0017	.0033	.0057	.0091	.0137		11
6		.0000	.0000	.0000	.0000	.0001	.0002	.0005	.0009	.0017	.0028		10
7		.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0002	.0004		9
8		.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001		8
		.99	.98	.97	.96	.95	.94	.93	.92	.91	.90	P	R

R	P	.11	.12	.13	.14	.15	.16	.17	.18	.19	.20		
0		.1550	.1293	.1077	.0895	.0743	.0614	.0507	.0418	.0343	.0281		16
1		.3065	.2822	.2575	.2332	.2097	.1873	.1662	.1468	.1289	.1126		15
2		.2841	.2886	.2886	.2847	.2775	.2675	.2554	.2416	.2267	.2111		14
3		.1638	.1837	.2013	.2163	.2285	.2378	.2441	.2475	.2482	.2463		13
4		.0658	.0814	.0977	.1144	.1311	.1472	.1625	.1766	.1892	.2001		12
5		.0195	.0266	.0351	.0447	.0555	.0673	.0799	.0930	.1065	.1201		11
6		.0044	.0067	.0096	.0133	.0180	.0235	.0300	.0374	.0458	.0550		10
7		.0008	.0013	.0020	.0031	.0045	.0064	.0088	.0117	.0153	.0197		9
8		.0001	.0002	.0003	.0006	.0009	.0014	.0020	.0029	.0041	.0055		8
9		.0000	.0000	.0000	.0001	.0001	.0002	.0004	.0006	.0008	.0012		7
10		.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0001	.0001	.0002		6
		.89	.88	.87	.86	.85	.84	.83	.82	.81	.80	P	R

R	P	.21	.22	.23	.24	.25	.26	.27	.28	.29	.30		
0		.0230	.0188	.0153	.0124	.0100	.0081	.0065	.0052	.0042	.0033		16
1		.0979	.0847	.0730	.0626	.0535	.0455	.0385	.0325	.0273	.0228		15
2		.1952	.1792	.1635	.1482	.1336	.1198	.1068	.0947	.0835	.0732		14
3		.2421	.2359	.2279	.2185	.2079	.1964	.1843	.1718	.1591	.1465		13
4		.2092	.2162	.2212	.2242	.2252	.2243	.2215	.2171	.2112	.2040		12
5		.1334	.1464	.1586	.1699	.1802	.1891	.1966	.2026	.2071	.2099		11
6		.0650	.0757	.0869	.0984	.1101	.1218	.1333	.1445	.1551	.1649		10
7		.0247	.0305	.0371	.0444	.0524	.0611	.0704	.0803	.0905	.1010		9
8		.0074	.0097	.0125	.0158	.0197	.0242	.0293	.0351	.0416	.0487		8
9		.0017	.0024	.0033	.0044	.0058	.0075	.0096	.0121	.0151	.0185		7
10		.0003	.0005	.0007	.0010	.0014	.0019	.0025	.0033	.0043	.0056		6
11		.0000	.0001	.0001	.0002	.0002	.0004	.0005	.0007	.0010	.0013		5
12		.0000	.0000	.0000	.0000	.0000	.0001	.0001	.0001	.0002	.0002		4
		.79	.78	.77	.76	.75	.74	.73	.72	.71	.70	P	R

Table B-8 Binomial distribution (Continued)

R	P	.31	.32	.33	.34	.35	.36	.37	.38	.39	.40		
0		.0026	.0021	.0016	.0013	.0010	.0008	.0006	.0005	.0004	.0003		16
1		.0190	.0157	.0130	.0107	.0087	.0071	.0058	.0047	.0038	.0030		15
2		.0639	.0555	.0480	.0413	.0353	.0301	.0255	.0215	.0180	.0150		14
3		.1341	.1220	.1103	.0992	.0888	.0790	.0699	.0615	.0538	.0468		13
4		.1958	.1865	.1766	.1662	.1553	.1444	.1333	.1224	.1118	.1014		12
5		.2111	.2107	.2088	.2054	.2008	.1949	.1879	.1801	.1715	.1623		11
6		.1739	.1818	.1885	.1940	.1982	.2010	.2024	.2024	.2010	.1983		10
7		.1116	.1222	.1326	.1428	.1524	.1615	.1698	.1772	.1636	.1889		9
8		.0564	.0647	.0735	.0827	.0923	.1022	.1122	.1222	.1320	.1417		8
9		.0225	.0271	.0322	.0379	.0442	.0511	.0586	.0666	.0750	.0840		7
10		.0071	.0089	.0111	.0137	.0167	.0201	.0241	.0286	.0336	.0392		6
11		.0017	.0023	.0030	.0038	.0049	.0062	.0077	.0095	.0117	.0142		5
12		.0003	.0004	.0006	.0008	.0011	.0014	.0019	.0024	.0031	.0040		4
13		.0000	.0001	.0001	.0001	.0002	.0003	.0003	.0005	.0006	.0008		3
14		.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0001	.0001		2
		.69	.68	.67	.66	.65	.64	.63	.62	.61	.60	P	R

R	P	.41	.42	.43	.44	.45	.46	.47	.48	.49	.50		
0		.0002	.0002	.0001	.0001	.0001	.0001	.0000	.0000	.0000	.0000		16
1		.0024	.0019	.0015	.0012	.0009	.0007	.0006	.0004	.0003	.0002		15
2		.0125	.0103	.0085	.0069	.0056	.0046	.0037	.0029	.0023	.0018		14
3		.0405	.0349	.0299	.0254	.0215	.0181	.0151	.0126	.0104	.0085		13
4		.0915	.0821	.0732	.0649	.0572	.0501	.0436	.0378	.0325	.0278		12
5		.1526	.1426	.1325	.1224	.1123	.1024	.0929	.0837	.0749	.0667		11
6		.1944	.1894	.1833	.1762	.1684	.1600	.1510	.1416	.1319	.1222		10
7		.1930	.1959	.1975	.1978	.1969	.1947	.1912	.1867	.1811	.1746		9
8		.1509	.1596	.1676	.1749	.1812	.1865	.1908	.1939	.1958	.1964		8
9		.0932	.1027	.1124	.1221	.1318	.1413	.1504	.1591	.1672	.1746		7
10		.0453	.0521	.0594	.0672	.0755	.0842	.0934	.1028	.1124	.1222		6
11		.0172	.0206	.0244	.0288	.0337	.0391	.0452	.0518	.0589	.0667		5
12		.0050	.0062	.0077	.0094	.0115	.0139	.0167	.0199	.0236	.0278		4
13		.0011	.0014	.0018	.0023	.0029	.0036	.0046	.0057	.0070	.0085		3
14		.0002	.0002	.0003	.0004	.0005	.0007	.0009	.0011	.0014	.0018		2
15		.0000	.0000	.0000	.0000	.0001	.0001	.0001	.0001	.0002	.0002		1
		.59	.58	.57	.56	.55	.54	.53	.52	.51	.50	P	R

N = 17

R	P	.01	.02	.03	.04	.05	.06	.07	.08	.09	.10		
0		.8429	.7093	.5958	.4996	.4181	.3493	.2912	.2423	.2012	.1668		17
1		.1447	.2461	.3133	.3539	.3741	.3790	.3726	.3582	.3383	.3150		16
2		.0117	.0402	.0775	.1180	.1575	.1935	.2244	.2492	.2677	.2800		15
3		.0006	.0041	.0120	.0246	.0415	.0618	.0844	.1083	.1324	.1556		14
4		.0000	.0003	.0013	.0036	.0075	.0138	.0222	.0330	.0458	.0605		13
5		.0000	.0000	.0001	.0004	.0010	.0023	.0044	.0075	.0118	.0175		12
6		.0000	.0000	.0000	.0000	.0001	.0003	.0007	.0013	.0023	.0039		11
7		.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0002	.0004	.0007		10
8		.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001		9
		.99	.98	.97	.96	.95	.94	.93	.92	.91	.90	P	R

R	P	.11	.12	.13	.14	.15	.16	.17	.18	.19	.20		
0		.1379	.1138	.0937	.0770	.0631	.0516	.0421	.0343	.0278	.0225		17
1		.2898	.2638	.2381	.2131	.1893	.1671	.1466	.1279	.1109	.0957		16
2		.2865	.2878	.2846	.2775	.2673	.2547	.2402	.2245	.2081	.1914		15
3		.1771	.1963	.2126	.2259	.2359	.2425	.2460	.2464	.2441	.2393		14
4		.0766	.0937	.1112	.1287	.1457	.1617	.1764	.1893	.2004	.2093		13
5		.0246	.0332	.0432	.0545	.0668	.0801	.0939	.1081	.1222	.1361		12
6		.0061	.0091	.0129	.0177	.0236	.0305	.0385	.0474	.0573	.0680		11
7		.0012	.0019	.0030	.0045	.0065	.0091	.0124	.0164	.0211	.0267		10
8		.0002	.0003	.0006	.0009	.0014	.0022	.0032	.0045	.0062	.0084		9
9		.0000	.0000	.0001	.0002	.0003	.0004	.0006	.0010	.0015	.0021		8
10		.0000	.0000	.0000	.0000	.0000	.0001	.0001	.0002	.0003	.0004		7
11		.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001		6
		.89	.88	.87	.86	.85	.84	.83	.82	.81	.80	P	R

Table B-8 Binomial distribution (Continued)

R	P	.21	.22	.23	.24	.25	.26	.27	.28	.29	.30		
0		.0182	.0146	.0118	.0094	.0075	.0060	.0047	.0038	.0030	.0023		17
1		.0822	.0702	.0597	.0505	.0426	.0357	.0299	.0248	.0206	.0169		16
2		.1747	.1584	.1427	.1277	.1136	.1005	.0883	.0772	.0672	.0581		15
3		.2322	.2234	.2131	.2016	.1893	.1765	.1634	.1502	.1372	.1245		14
4		.2161	.2205	.2228	.2228	.2209	.2170	.2115	.2044	.1961	.1868		13
5		.1493	.1617	.1730	.1830	.1914	.1982	.2033	.2067	.2083	.2081		12
6		.0794	.0912	.1034	.1156	.1276	.1393	.1504	.1608	.1701	.1784		11
7		.0332	.0404	.0485	.0573	.0668	.0769	.0874	.0982	.1092	.1201		10
8		.0110	.0143	.0181	.0226	.0279	.0338	.0404	.0478	.0558	.0644		9
9		.0029	.0040	.0054	.0071	.0093	.0119	.0150	.0186	.0228	.0276		8
10		.0006	.0009	.0013	.0018	.0025	.0033	.0044	.0058	.0074	.0095		7
11		.0001	.0002	.0002	.0004	.0005	.0007	.0010	.0014	.0019	.0026		6
12		.0000	.0000	.0000	.0001	.0001	.0001	.0002	.0003	.0004	.0006		5
13		.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0001		4
		.79	.78	.77	.76	.75	.74	.73	.72	.71	.70	P	R

R	P	.31	.32	.33	.34	.35	.36	.37	.38	.39	.40		
0		.0018	.0014	.0011	.0009	.0007	.0005	.0004	.0003	.0002	.0002		17
1		.0139	.0114	.0093	.0075	.0060	.0048	.0039	.0031	.0024	.0019		16
2		.0500	.0428	.0364	.0309	.0260	.0218	.0182	.0151	.0125	.0102		15
3		.1123	.1007	.0898	.0795	.0701	.0614	.0534	.0463	.0398	.0341		14
4		.1766	.1659	.1547	.1434	.1320	.1208	.1099	.0993	.0892	.0796		13
5		.2063	.2030	.1982	.1921	.1849	.1767	.1677	.1582	.1482	.1379		12
6		.1854	.1910	.1952	.1979	.1991	.1988	.1970	.1939	.1895	.1839		11
7		.1309	.1413	.1511	.1602	.1685	.1757	.1818	.1868	.1904	.1927		10
8		.0735	.0831	.0930	.1032	.1134	.1235	.1335	.1431	.1521	.1606		9
9		.0330	.0391	.0458	.0531	.0611	.0695	.0784	.0877	.0973	.1070		8
10		.0119	.0147	.0181	.0219	.0263	.0313	.0368	.0430	.0498	.0571		7
11		.0034	.0044	.0057	.0072	.0090	.0112	.0138	.0168	.0202	.0242		6
12		.0008	.0010	.0014	.0018	.0024	.0031	.0040	.0051	.0065	.0081		5
13		.0001	.0002	.0003	.0004	.0005	.0007	.0009	.0012	.0016	.0021		4
14		.0000	.0000	.0000	.0001	.0001	.0001	.0002	.0002	.0003	.0004		3
15		.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001		2
		.69	.68	.67	.66	.65	.64	.63	.62	.61	.60	P	R

R	P	.41	.42	.43	.44	.45	.46	.47	.48	.49	.50		
0		.0001	.0001	.0001	.0001	.0000	.0000	.0000	.0000	.0000	.0000		17
1		.0015	.0012	.0009	.0007	.0005	.0004	.0003	.0002	.0002	.0001		16
2		.0084	.0068	.0055	.0044	.0035	.0028	.0022	.0017	.0013	.0010		15
3		.0290	.0246	.0207	.0173	.0144	.0119	.0097	.0079	.0064	.0052		14
4		.0706	.0622	.0546	.0475	.0411	.0354	.0302	.0257	.0217	.0182		13
5		.1276	.1172	.1070	.0971	.0875	.0784	.0697	.0616	.0541	.0472		12
6		.1773	.1697	.1614	.1525	.1432	.1335	.1237	.1138	.1040	.0944		11
7		.1936	.1932	.1914	.1883	.1841	.1787	.1723	.1650	.1570	.1484		10
8		.1682	.1748	.1805	.1850	.1883	.1903	.1910	.1904	.1886	.1855		9
9		.1169	.1266	.1361	.1453	.1540	.1621	.1694	.1758	.1812	.1855		8
10		.0650	.0733	.0822	.0914	.1008	.1105	.1202	.1298	.1393	.1484		7
11		.0287	.0338	.0394	.0457	.0525	.0599	.0678	.0763	.0851	.0944		6
12		.0100	.0122	.0149	.0179	.0215	.0255	.0301	.0352	.0409	.0472		5
13		.0027	.0034	.0043	.0054	.0068	.0084	.0103	.0125	.0151	.0182		4
14		.0005	.0007	.0009	.0012	.0015	.0020	.0026	.0033	.0041	.0052		3
15		.0001	.0001	.0001	.0002	.0003	.0003	.0005	.0006	.0008	.0010		2
16		.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0001	.0001	.0001		1
		.59	.58	.57	.56	.55	.54	.53	.52	.51	.50	P	R

18

R	P	.01	.02	.03	.04	.05	.06	.07	.08	.09	.10	
0		.8345	.6951	.5780	.4796	.3972	.3283	.2708	.2229	.1831	.1501	18

Table B-8 Binomial distribution (Continued)

1		.1517	.2554	.3217	.3597	.3763	.3772	.3669	.3489	.3260	.3002	17
2		.0130	.0443	.0846	.1274	.1683	.2047	.2348	.2579	.2741	.2835	16
3		.0007	.0048	.0140	.0283	.0473	.0697	.0942	.1196	.1446	.1680	15
4		.0000	.0004	.0016	.0044	.0093	.0167	.0266	.0390	.0536	.0700	14
5		.0000	.0000	.0001	.0005	.0014	.0030	.0056	.0095	.0148	.0218	13
6		.0000	.0000	.0000	.0000	.0002	.0004	.0009	.0018	.0032	.0052	12
7		.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0003	.0005	.0010	11
8		.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0002	10
		.99	.98	.97	.96	.95	.94	.93	.92	.91	.90	P R

R	P	.11	.12	.13	.14	.15	.16	.17	.18	.19	.20	
0		.1227	.1002	.0815	.0662	.0536	.0434	.0349	.0281	.0225	.0180	18
1		.2731	.2458	.2193	.1940	.1704	.1486	.1288	.1110	.0951	.0811	17
2		.2869	.2850	.2785	.2685	.2556	.2407	.2243	.2071	.1897	.1723	16
3		.1891	.2072	.2220	.2331	.2406	.2445	.2450	.2425	.2373	.2297	15
4		.0877	.1060	.1244	.1423	.1592	.1746	.1882	.1996	.2087	.2153	14
5		.0303	.0405	.0520	.0649	.0787	.0931	.1079	.1227	.1371	.1507	13
6		.0081	.0120	.0168	.0229	.0301	.0384	.0479	.0584	.0697	.0816	12
7		.0017	.0028	.0043	.0064	.0091	.0126	.0168	.0220	.0280	.0350	11
8		.0003	.0005	.0009	.0014	.0022	.0033	.0047	.0066	.0090	.0120	10
9		.0000	.0001	.0001	.0003	.0004	.0007	.0011	.0016	.0024	.0033	9
10		.0000	.0000	.0000	.0000	.0001	.0001	.0002	.0003	.0005	.0008	8
11		.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0001	.0001	7
		.89	.88	.87	.86	.85	.84	.83	.82	.81	.80	P R

R	P	.21	.22	.23	.24	.25	.26	.27	.28	.29	.30	
0		.0144	.0114	.0091	.0072	.0056	.0044	.0035	.0027	.0021	.0016	18
1		.0687	.0580	.0487	.0407	.0338	.0280	.0231	.0189	.0155	.0126	17
2		.1553	.1390	.1236	.1092	.0958	.0836	.0725	.0626	.0537	.0458	16
3		.2202	.2091	.1969	.1839	.1704	.1567	.1431	.1298	.1169	.1046	15
4		.2195	.2212	.2205	.2177	.2130	.2065	.1985	.1892	.1790	.1681	14
5		.1634	.1747	.1845	.1925	.1988	.2031	.2055	.2061	.2048	.2017	13
6		.0941	.1067	.1194	.1317	.1436	.1546	.1647	.1736	.1812	.1873	12
7		.0429	.0516	.0611	.0713	.0820	.0931	.1044	.1157	.1269	.1376	11
8		.0157	.0200	.0251	.0310	.0376	.0450	.0531	.0619	.0713	.0811	10
9		.0046	.0063	.0083	.0109	.0139	.0176	.0218	.0267	.0323	.0386	9
10		.0011	.0016	.0022	.0031	.0042	.0056	.0073	.0094	.0119	.0149	8
11		.0002	.0003	.0005	.0007	.0010	.0014	.0020	.0026	.0035	.0046	7
12		.0000	.0001	.0001	.0001	.0002	.0003	.0004	.0006	.0008	.0012	6
13		.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0001	.0002	.0002	5
		.79	.78	.77	.76	.75	.74	.73	.72	.71	.70	P R

R	P	.31	.32	.33	.34	.35	.36	.37	.38	.39	.40	
0		.0013	.0010	.0007	.0006	.0004	.0003	.0002	.0002	.0001	.0001	18
1		.0102	.0082	.0066	.0052	.0042	.0033	.0026	.0020	.0016	.0012	17
2		.0388	.0327	.0275	.0229	.0190	.0157	.0129	.0105	.0086	.0069	16
3		.0930	.0822	.0722	.0630	.0547	.0471	.0404	.0344	.0292	.0246	15
4		.1567	.1450	.1333	.1217	.1104	.0994	.0890	.0791	.0699	.0614	14
5		.1971	.1911	.1838	.1755	.1664	.1566	.1463	.1358	.1252	.1146	13
6		.1919	.1948	.1962	.1959	.1941	.1908	.1862	.1803	.1734	.1655	12
7		.1478	.1572	.1656	.1730	.1792	.1840	.1875	.1895	.1900	.1892	11
8		.0913	.1017	.1122	.1226	.1327	.1423	.1514	.1597	.1671	.1734	10
9		.0456	.0532	.0614	.0701	.0794	.0890	.0988	.1087	.1187	.1284	9
10		.0184	.0225	.0272	.0325	.0385	.0450	.0522	.0600	.0683	.0771	8
11		.0060	.0077	.0097	.0122	.0151	.0184	.0223	.0267	.0318	.0374	7
12		.0016	.0021	.0028	.0037	.0047	.0060	.0076	.0096	.0118	.0145	6
13		.0003	.0005	.0006	.0009	.0012	.0016	.0021	.0027	.0035	.0045	5
14		.0001	.0001	.0001	.0002	.0002	.0003	.0004	.0006	.0008	.0011	4
15		.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0001	.0001	.0002	3
		.69	.68	.67	.66	.65	.64	.63	.62	.61	.60	P R

Table B-8 Binomial distribution (Continued)

R	P	.41	.42	.43	.44	.45	.46	.47	.48	.49	.50		
0		.0001	.0001	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000		18
1		.0009	.0007	.0005	.0004	.0003	.0002	.0002	.0001	.0001	.0001		17
2		.0055	.0044	.0035	.0028	.0022	.0017	.0013	.0010	.0008	.0006		16
3		.0206	.0171	.0141	.0116	.0095	.0077	.0062	.0050	.0039	.0031		15
4		.0536	.0464	.0400	.0342	.0291	.0246	.0206	.0172	.0142	.0117		14
5		.1042	.0941	.0844	.0753	.0666	.0586	.0512	.0444	.0382	.0327		13
6		.1569	.1477	.1380	.1281	.1181	.1081	.0983	.0887	.0796	.0708		12
7		.1869	.1833	.1785	.1726	.1657	.1579	.1494	.1404	.1310	.1214		11
8		.1786	.1825	.1852	.1864	.1864	.1850	.1822	.1782	.1731	.1669		10
9		.1379	.1469	.1552	.1628	.1694	.1751	.1795	.1828	.1848	.1855		9
10		.0862	.0957	.1054	.1151	.1248	.1342	.1433	.1519	.1598	.1669		8
11		.0436	.0504	.0578	.0658	.0742	.0831	.0924	.1020	.1117	.1214		7
12		.0177	.0213	.0254	.0301	.0354	.0413	.0478	.0549	.0626	.0708		6
13		.0057	.0071	.0089	.0109	.0134	.0162	.0196	.0234	.0278	.0327		5
14		.0014	.0018	.0024	.0031	.0039	.0049	.0062	.0077	.0095	.0117		4
15		.0003	.0004	.0005	.0006	.0009	.0011	.0015	.0019	.0024	.0031		3
16		.0000	.0000	.0001	.0001	.0001	.0002	.0002	.0003	.0004	.0006		2
17		.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001		1
		.59	.58	.57	.56	.55	.54	.53	.52	.51	.50	P	R

N = 19

R	P	.01	.02	.03	.04	.05	.06	.07	.08	.09	.10		
0		.8262	.6812	.5606	.4604	.3774	.3086	.2519	.2051	.1666	.1351		19
1		.1586	.2642	.3294	.3645	.3774	.3743	.3602	.3389	.3131	.2852		18
2		.0144	.0485	.0917	.1367	.1787	.2150	.2440	.2652	.2787	.2852		17
3		.0008	.0056	.0161	.0323	.0533	.0778	.1041	.1307	.1562	.1796		16
4		.0000	.0005	.0020	.0054	.0112	.0199	.0313	.0455	.0618	.0798		15
5		.0000	.0000	.0002	.0007	.0018	.0038	.0071	.0119	.0183	.0266		14
6		.0000	.0000	.0000	.0001	.0002	.0006	.0012	.0024	.0042	.0069		13
7		.0000	.0000	.0000	.0000	.0000	.0001	.0002	.0004	.0008	.0014		12
8		.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0001	.0002		11
		.99	.98	.97	.96	.95	.94	.93	.92	.91	.90	P	R

R	P	.11	.12	.13	.14	.15	.16	.17	.18	.19	.20		
0		.1092	.0881	.0709	.0569	.0456	.0364	.0290	.0230	.0182	.0144		19
1		.2565	.2284	.2014	.1761	.1529	.1318	.1129	.0961	.0813	.0685		18
2		.2854	.2803	.2708	.2581	.2428	.2259	.2081	.1898	.1717	.1540		17
3		.1999	.2166	.2293	.2381	.2428	.2439	.2415	.2361	.2282	.2182		16
4		.0988	.1181	.1371	.1550	.1714	.1858	.1979	.2073	.2141	.2182		15
5		.0366	.0483	.0614	.0757	.0907	.1062	.1216	.1365	.1507	.1636		14
6		.0106	.0154	.0214	.0288	.0374	.0472	.0581	.0699	.0825	.0955		13
7		.0024	.0039	.0059	.0087	.0122	.0167	.0221	.0285	.0359	.0443		12
8		.0004	.0008	.0013	.0021	.0032	.0048	.0068	.0094	.0126	.0166		11
9		.0001	.0001	.0002	.0004	.0007	.0011	.0017	.0025	.0036	.0051		10
10		.0000	.0000	.0000	.0001	.0001	.0002	.0003	.0006	.0009	.0013		9
11		.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0001	.0002	.0003		8
		.89	.88	.87	.86	.85	.84	.83	.82	.81	.80	P	R

R	P	.21	.22	.23	.24	.25	.26	.27	.28	.29	.30		
0		.0113	.0089	.0070	.0054	.0042	.0033	.0025	.0019	.0015	.0011		19
1		.0573	.0477	.0396	.0326	.0268	.0219	.0178	.0144	.0116	.0093		18
2		.1371	.1212	.1064	.0927	.0803	.0692	.0592	.0503	.0426	.0358		17
3		.2065	.1937	.1800	.1659	.1517	.1377	.1240	.1109	.0985	.0869		16
4		.2196	.2185	.2151	.2096	.2023	.1935	.1835	.1726	.1610	.1491		15
5		.1751	.1849	.1928	.1986	.2023	.2040	.2036	.2013	.1973	.1916		14
6		.1086	.1217	.1343	.1463	.1574	.1672	.1757	.1827	.1880	.1916		13
7		.0536	.0637	.0745	.0858	.0974	.1091	.1207	.1320	.1426	.1525		12
8		.0214	.0270	.0334	.0406	.0487	.0575	.0670	.0770	.0874	.0981		11
9		.0069	.0093	.0122	.0157	.0198	.0247	.0303	.0366	.0436	.0514		10
10		.0018	.0026	.0036	.0050	.0066	.0087	.0112	.0142	.0178	.0220		9
11		.0004	.0006	.0009	.0013	.0018	.0025	.0034	.0045	.0060	.0077		8
12		.0001	.0001	.0002	.0003	.0004	.0005	.0008	.0012	.0016	.0022		7
13		.0000	.0000	.0000	.0000	.0001	.0001	.0002	.0002	.0004	.0005		6
14		.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0001		5
		.79	.78	.77	.76	.75	.74	.73	.72	.71	.70	P	R

Table B-8 Binomial distribution (Continued)

R	P	.31	.32	.33	.34	.35	.36	.37	.38	.39	.40		
0		.0009	.0007	.0005	.0004	.0003	.0002	.0002	.0001	.0001	.0001		19
1		.0074	.0059	.0046	.0036	.0029	.0022	.0017	.0013	.0010	.0008		18
2		.0299	.0249	.0206	.0169	.0138	.0112	.0091	.0073	.0058	.0046		17
3		.0762	.0664	.0574	.0494	.0422	.0358	.0302	.0253	.0211	.0175		16
4		.1370	.1249	.1131	.1017	.0909	.0806	.0710	.0621	.0540	.0467		15
5		.1846	.1764	.1672	.1572	.1468	.1360	.1251	.1143	.1036	.0933		14
6		.1935	.1936	.1921	.1890	.1844	.1785	.1714	.1634	.1546	.1451		13
7		.1615	.1692	.1757	.1808	.1844	.1865	.1870	.1860	.1835	.1797		12
8		.1088	.1195	.1298	.1397	.1489	.1573	.1647	.1710	.1760	.1797		11
9		.0597	.0687	.0782	.0880	.0980	.1082	.1182	.1281	.1375	.1464		10
10		.0268	.0323	.0385	.0453	.0528	.0608	.0694	.0785	.0879	.0976		9
11		.0099	.0124	.0155	.0191	.0233	.0280	.0334	.0394	.0460	.0532		8
12		.0030	.0039	.0051	.0066	.0083	.0105	.0131	.0161	.0196	.0237		7
13		.0007	.0010	.0014	.0018	.0024	.0032	.0041	.0053	.0068	.0085		6
14		.0001	.0002	.0003	.0004	.0006	.0008	.0010	.0014	.0018	.0024		5
15		.0000	.0000	.0000	.0001	.0001	.0001	.0002	.0003	.0004	.0005		4
16		.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0001		3
		.69	.68	.67	.66	.65	.64	.63	.62	.61	.60	P	R

R	P	.41	.42	.43	.44	.45	.46	.47	.48	.49	.50		
1		.0006	.0004	.0003	.0002	.0002	.0001	.0001	.0001	.0001	.0000		18
2		.0037	.0029	.0022	.0017	.0013	.0010	.0008	.0006	.0004	.0003		17
3		.0144	.0118	.0096	.0077	.0062	.0049	.0039	.0031	.0024	.0018		16
4		.0400	.0341	.0289	.0243	.0203	.0165	.0138	.0113	.0092	.0074		15
5		.0834	.0741	.0653	.0572	.0497	.0429	.0368	.0313	.0265	.0222		14
6		.1353	.1252	.1150	.1049	.0949	.0853	.0761	.0674	.0593	.0518		13
7		.1746	.1683	.1611	.1530	.1443	.1350	.1254	.1156	.1058	.0961		12
8		.1820	.1829	.1823	.1803	.1771	.1725	.1668	.1601	.1525	.1442		11
9		.1546	.1618	.1681	.1732	.1771	.1796	.1808	.1806	.1791	.1762		10
10		.1074	.1172	.1268	.1361	.1449	.1530	.1603	.1667	.1721	.1762		9
11		.0611	.0694	.0783	.0875	.0970	.1066	.1163	.1259	.1352	.1442		8
12		.0283	.0335	.0394	.0458	.0529	.0606	.0688	.0775	.0866	.0961		7
13		.0106	.0131	.0160	.0194	.0233	.0278	.0328	.0385	.0448	.0518		6
14		.0032	.0041	.0052	.0065	.0082	.0101	.0125	.0152	.0185	.0222		5
15		.0007	.0010	.0013	.0017	.0022	.0029	.0037	.0047	.0059	.0074		4
16		.0001	.0002	.0002	.0003	.0005	.0006	.0008	.0011	.0014	.0018		3
17		.0000	.0000	.0000	.0000	.0001	.0001	.0001	.0002	.0002	.0003		2
		.59	.58	.57	.56	.55	.54	.53	.52	.51	.50	P	R

N = 20

R	P	.01	.02	.03	.04	.05	.06	.07	.08	.09	.10		
0		.8179	.6676	.5438	.4420	.3585	.2901	.2342	.1887	.1516	.1216		20
1		.1652	.2725	.3364	.3683	.3774	.3703	.3526	.3282	.3000	.2702		19
2		.0159	.0528	.0988	.1458	.1887	.2246	.2521	.2711	.2818	.2852		18
3		.0010	.0065	.0183	.0364	.0596	.0860	.1139	.1414	.1672	.1901		17
4		.0000	.0006	.0024	.0065	.0133	.0233	.0364	.0523	.0703	.0898		16
5		.0000	.0000	.0002	.0009	.0022	.0048	.0088	.0145	.0222	.0319		15
6		.0000	.0000	.0000	.0001	.0003	.0008	.0017	.0032	.0055	.0089		14
7		.0000	.0000	.0000	.0000	.0000	.0001	.0002	.0006	.0011	.0020		13
8		.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0002	.0004		12
9		.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001		11
		.99	.98	.97	.96	.95	.94	.93	.92	.91	.90	P	R

R	P	.11	.12	.13	.14	.15	.16	.17	.18	.19	.20		
0		.0972	.0776	.0617	.0490	.0388	.0306	.0241	.0189	.0148	.0115		20
1		.2403	.2115	.1844	.1595	.1368	.1165	.0986	.0829	.0693	.0576		19
2		.2822	.2740	.2618	.2466	.2293	.2109	.1919	.1730	.1545	.1369		18
3		.2093	.2242	.2347	.2409	.2428	.2410	.2358	.2278	.2175	.2054		17
4		.1099	.1299	.1491	.1666	.1821	.1951	.2053	.2125	.2168	.2182		16
5		.0435	.0567	.0713	.0868	.1028	.1189	.1345	.1493	.1627	.1746		15
6		.0134	.0193	.0266	.0353	.0454	.0566	.0689	.0819	.0954	.1091		14
7		.0033	.0053	.0080	.0115	.0160	.0216	.0282	.0360	.0448	.0545		13
8		.0007	.0012	.0019	.0030	.0046	.0067	.0094	.0128	.0171	.0222		12
9		.0001	.0002	.0004	.0007	.0011	.0017	.0026	.0038	.0053	.0074		11
10		.0000	.0000	.0001	.0001	.0002	.0004	.0006	.0009	.0014	.0020		10

Table B-8 Binomial distribution (Continued)

11		.0000	.0000	.0000	.0000	.0000	.0001	.0001	.0002	.0003	.0005		9
12		.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0001		8
		.89	.88	.87	.86	.85	.84	.83	.82	.81	.80	P	R

R	P	.21	.22	.23	.24	.25	.26	.27	.28	.29	.30		
0		.0090	.0069	.0054	.0041	.0032	.0024	.0018	.0014	.0011	.0008		20
1		.0477	.0392	.0321	.0261	.0211	.0170	.0137	.0109	.0087	.0068		19
2		.1204	.1050	.0910	.0783	.0669	.0569	.0480	.0403	.0336	.0278		18
3		.1920	.1777	.1631	.1484	.1339	.1199	.1065	.0940	.0823	.0716		17
4		.2169	.2131	.2070	.1991	.1897	.1790	.1675	.1553	.1429	.1304		16
5		.1845	.1923	.1979	.2012	.2023	.2013	.1982	.1933	.1868	.1789		15
6		.1226	.1356	.1478	.1589	.1686	.1768	.1833	.1879	.1907	.1916		14
7		.0652	.0765	.0883	.1003	.1124	.1242	.1356	.1462	.1558	.1643		13
8		.0282	.0351	.0429	.0515	.0609	.0709	.0815	.0924	.1034	.1144		12
9		.0100	.0132	.0171	.0217	.0271	.0332	.0402	.0479	.0563	.0654		11
10		.0029	.0041	.0056	.0075	.0099	.0128	.0163	.0205	.0253	.0308		10
11		.0007	.0010	.0015	.0022	.0030	.0041	.0055	.0072	.0094	.0120		9
12		.0001	.0002	.0003	.0005	.0008	.0011	.0015	.0021	.0029	.0039		8
13		.0000	.0000	.0001	.0001	.0002	.0002	.0003	.0005	.0007	.0010		7
14		.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0001	.0001	.0002		6
		.79	.78	.77	.76	.75	.74	.73	.72	.71	.70	P	R

R	P	.31	.32	.33	.34	.35	.36	.37	.38	.39	.40		
0		.0006	.0004	.0003	.0002	.0002	.0001	.0001	.0001	.0001	.0000		20
1		.0054	.0042	.0033	.0025	.0020	.0015	.0011	.0009	.0007	.0005		19
2		.0229	.0188	.0153	.0124	.0100	.0080	.0064	.0050	.0040	.0031		18
3		.0619	.0531	.0453	.0383	.0323	.0270	.0224	.0185	.0152	.0123		17
4		.1181	.1062	.0947	.0839	.0738	.0645	.0559	.0482	.0412	.0350		16
5		.1698	.1599	.1493	.1384	.1272	.1161	.1051	.0945	.0843	.0746		15
6		.1907	.1881	.1839	.1782	.1712	.1632	.1543	.1447	.1347	.1244		14
7		.1714	.1770	.1811	.1836	.1844	.1836	.1812	.1774	.1722	.1659		13
8		.1251	.1354	.1450	.1537	.1614	.1678	.1730	.1767	.1790	.1797		12
9		.0750	.0849	.0952	.1056	.1158	.1259	.1354	.1444	.1526	.1597		11
10		.0370	.0440	.0516	.0598	.0686	.0779	.0875	.0974	.1073	.1171		10
11		.0151	.0188	.0231	.0280	.0336	.0398	.0467	.0542	.0624	.0710		9
12		.0051	.0066	.0085	.0108	.0136	.0168	.0206	.0249	.0299	.0355		8
13		.0014	.0019	.0026	.0034	.0045	.0058	.0074	.0094	.0118	.0146		7
14		.0003	.0005	.0006	.0009	.0012	.0016	.0022	.0029	.0038	.0049		6
15		.0001	.0001	.0001	.0002	.0003	.0004	.0005	.0007	.0010	.0013		5
16		.0000	.0000	.0000	.0000	.0000	.0001	.0001	.0001	.0002	.0003		4
		.69	.68	.67	.66	.65	.64	.63	.62	.61	.60	P	R

R	P	.41	.42	.43	.44	.45	.46	.47	.48	.49	.50		
1		.0004	.0003	.0002	.0001	.0001	.0001	.0001	.0000	.0000	.0000		19
2		.0024	.0018	.0014	.0011	.0008	.0006	.0005	.0003	.0002	.0002		18
3		.0100	.0080	.0064	.0051	.0040	.0031	.0024	.0019	.0014	.0011		17
4		.0295	.0247	.0206	.0170	.0139	.0113	.0092	.0074	.0059	.0046		16
5		.0656	.0573	.0496	.0427	.0365	.0309	.0260	.0217	.0180	.0148		15
6		.1140	.1037	.0936	.0839	.0746	.0658	.0577	.0501	.0432	.0370		14
7		.1585	.1502	.1413	.1318	.1221	.1122	.1023	.0925	.0830	.0739		13
8		.1790	.1768	.1732	.1683	.1623	.1553	.1474	.1388	.1296	.1201		12
9		.1658	.1707	.1742	.1763	.1771	.1763	.1742	.1708	.1661	.1602		11
10		.1268	.1359	.1446	.1524	.1593	.1652	.1700	.1734	.1755	.1762		10
11		.0801	.0895	.0991	.1089	.1185	.1280	.1370	.1455	.1533	.1602		9
12		.0417	.0486	.0561	.0642	.0727	.0818	.0911	.1007	.1105	.1201		8
13		.0178	.0217	.0260	.0310	.0366	.0429	.0497	.0572	.0653	.0739		7
14		.0062	.0078	.0098	.0122	.0150	.0183	.0221	.0264	.0314	.0370		6
15		.0017	.0023	.0030	.0038	.0049	.0062	.0078	.0098	.0121	.0148		5
16		.0004	.0005	.0007	.0009	.0013	.0017	.0022	.0028	.0036	.0046		4
17		.0001	.0001	.0001	.0002	.0002	.0003	.0005	.0006	.0008	.0011		3
18		.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0001	.0001	.0002		2
		.59	.58	.57	.56	.55	.54	.53	.52	.51	.50	P	R

Table B-8 Binomial distribution (Continued)

N = 50

R	P	.01	.02	.03	.04	.05	.06	.07	.08	.09	.10		
0		.6050	.3642	.2181	.1299	.0769	.0453	.0266	.0155	.0090	.0052		50
1		.3056	.3716	.3372	.2706	.2025	.1447	.0999	.0672	.0443	.0286		49
2		.0756	.1858	.2555	.2762	.2611	.2262	.1843	.1433	.1073	.0779		48
3		.0122	.0607	.1264	.1842	.2199	.2311	.2219	.1993	.1698	.1386		47
4		.0015	.0145	.0459	.0902	.1360	.1733	.1963	.2037	.1973	.1809		46
5		.0001	.0027	.0131	.0346	.0658	.1018	.1359	.1629	.1795	.1849		45
6		.0000	.0004	.0030	.0108	.0260	.0487	.0767	.1063	.1332	.1541		44
7		.0000	.0001	.0006	.0028	.0085	.0195	.0363	.0581	.0828	.1076		43
8		.0000	.0000	.0001	.0006	.0024	.0067	.0147	.0271	.0440	.0643		42
9		.0000	.0000	.0000	.0001	.0006	.0020	.0052	.0110	.0203	.0333		41
10		.0000	.0000	.0000	.0000	.0001	.0005	.0016	.0039	.0082	.0152		40
11		.0000	.0000	.0000	.0000	.0000	.0001	.0004	.0012	.0030	.0061		39
12		.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0004	.0010	.0022		38
13		.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0003	.0007		37
14		0	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0002		36
15		0	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001		35
		.99	.98	.97	.96	.95	.94	.93	.92	.91	.90	P	R

R	P	.11	.12	.13	.14	.15	.16	.17	.18	.19	.20		
0		.0029	.0017	.0009	.0005	.0003	.0002	.0001	.0000	.0000	.0000		50
1		.0182	.0114	.0071	.0043	.0026	.0016	.0009	.0005	.0003	.0002		49
2		.0552	.0382	.0259	.0172	.0113	.0073	.0046	.0029	.0018	.0011		48
3		.1091	.0833	.0619	.0449	.0319	.0222	.0151	.0102	.0067	.0044		47
4		.1584	.1334	.1086	.0858	.0661	.0496	.0364	.0262	.0185	.0128		46
5		.1801	.1674	.1493	.1286	.1072	.0869	.0687	.0530	.0400	.0295		45
6		.1670	.1712	.1674	.1570	.1419	.1242	.1055	.0872	.0703	.0554		44
7		.1297	.1467	.1572	.1606	.1575	.1487	.1358	.1203	.1037	.0870		43
8		.0862	.1075	.1263	.1406	.1493	.1523	.1495	.1420	.1307	.1169		42
9		.0497	.0684	.0880	.1068	.1230	.1353	.1429	.1454	.1431	.1364		41
10		.0252	.0383	.0539	.0713	.0890	.1057	.1200	.1309	.1376	.1398		40
11		.0113	.0190	.0293	.0422	.0571	.0732	.0894	.1045	.1174	.1271		39
12		.0045	.0084	.0142	.0223	.0328	.0453	.0595	.0745	.0895	.1033		38
13		.0016	.0034	.0062	.0106	.0169	.0252	.0356	.0478	.0613	.0755		37
14		.0005	.0012	.0025	.0046	.0079	.0127	.0193	.0277	.0380	.0499		36
15		.0002	.0004	.0009	.0018	.0033	.0058	.0095	.0146	.0214	.0299		35
16		.0000	.0001	.0003	.0006	.0013	.0024	.0042	.0070	.0110	.0164		34
17		.0000	.0000	.0001	.0002	.0005	.0009	.0017	.0031	.0052	.0082		33
18		.0000	.0000	.0000	.0001	.0001	.0003	.0007	.0012	.0022	.0037		32
19		.0000	.0000	.0000	.0000	.0000	.0001	.0002	.0005	.0009	.0016		31
20		.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0002	.0003	.0006		30
21		.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0002		29
22		.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001		28
		.89	.88	.87	.86	.85	.84	.83	.82	.81	.80	P	R

R	P	.21	.22	.23	.24	.25	.26	.27	.28	.29	.30		
1		.0001	.0001	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000		49
2		.0007	.0004	.0002	.0001	.0001	.0000	.0000	.0000	.0000	.0000		48
3		.0028	.0018	.0011	.0007	.0004	.0002	.0001	.0001	.0000	.0000		47
4		.0088	.0059	.0039	.0025	.0015	.0010	.0006	.0004	.0002	.0001		46
5		.0214	.0152	.0106	.0073	.0049	.0033	.0021	.0014	.0009	.0006		45
6		.0427	.0322	.0238	.0173	.0123	.0087	.0060	.0040	.0027	.0018		44
7		.0713	.0571	.0447	.0344	.0259	.0191	.0139	.0099	.0069	.0048		43
8		.1019	.0865	.0718	.0583	.0463	.0361	.0276	.0207	.0152	.0110		42
9		.1263	.1139	.1001	.0859	.0721	.0592	.0476	.0375	.0290	.0220		41
10		.1377	.1317	.1226	.1113	.0985	.0852	.0721	.0598	.0485	.0386		40
11		.1331	.1351	.1332	.1278	.1194	.1089	.0970	.0845	.0721	.0602		39
12		.1150	.1238	.1293	.1311	.1294	.1244	.1166	.1068	.0957	.0838		38
13		.0894	.1021	.1129	.1210	.1261	.1277	.1261	.1215	.1142	.1050		37
14		.0628	.0761	.0891	.1010	.1110	.1186	.1233	.1248	.1233	.1189		36
15		.0400	.0515	.0639	.0766	.0888	.1000	.1094	.1165	.1209	.1223		35
16		.0233	.0318	.0417	.0529	.0648	.0769	.0855	.0991	.1080	.1147		34
17		.0124	.0179	.0249	.0334	.0432	.0540	.0655	.0771	.0882	.0983		33
18		.0060	.0093	.0137	.0193	.0264	.0348	.0444	.0550	.0661	.0772		32
19		.0027	.0044	.0069	.0103	.0148	.0206	.0277	.0360	.0454	.0558		31

Table B-8 Binomial distribution (Continued)

20		.0011	.0019	.0032	.0050	.0077	.0112	.0159	.0217	.0288	.0370	30
21		.0004	.0008	.0014	.0023	.0035	.0056	.0084	.0121	.0168	.0227	29
22		.0001	.0003	.0005	.0009	.0016	.0026	.0041	.0062	.0090	.0128	28
23		.0000	.0001	.0002	.0004	.0005	.0011	.0018	.0029	.0045	.0067	27
24		.0000	.0000	.0001	.0001	.0002	.0004	.0008	.0013	.0021	.0032	26
25		.0000	.0000	.0000	.0000	.0001	.0002	.0003	.0005	.0009	.0014	25
26		.0000	.0000	.0000	.0000	.0000	.0001	.0001	.0002	.0003	.0006	24
27		.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0001	.0002	23
28		.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	22
		.79	.78	.77	.76	.75	.74	.73	.72	.71	.70	P R

R	P	.31	.32	.33	.34	.35	.36	.37	.38	.39	.40	
4		.0001	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	46
5		.0003	.0002	.0001	.0001	.0000	.0000	.0000	.0000	.0000	.0000	45
6		.0011	.0007	.0005	.0003	.0002	.0001	.0001	.0000	.0000	.0000	44
7		.0032	.0022	.0014	.0009	.0005	.0004	.0002	.0001	.0001	.0000	43
8		.0078	.0055	.0037	.0025	.0017	.0011	.0007	.0004	.0003	.0002	42
9		.0164	.0120	.0086	.0061	.0042	.0029	.0019	.0013	.0008	.0005	41
10		.0301	.0231	.0174	.0128	.0093	.0066	.0046	.0032	.0022	.0014	40
11		.0493	.0395	.0311	.0240	.0182	.0136	.0099	.0071	.0050	.0035	39
12		.0719	.0604	.0498	.0402	.0319	.0248	.0189	.0142	.0105	.0076	38
13		.0944	.0831	.0717	.0606	.0502	.0408	.0325	.0255	.0195	.0147	37
14		.1121	.1034	.0933	.0825	.0714	.0607	.0505	.0412	.0330	.0260	36
15		.1209	.1168	.1103	.1020	.0923	.0819	.0712	.0606	.0507	.0415	35
16		.1188	.1202	.1189	.1149	.1088	.1008	.0914	.0813	.0709	.0606	34
17		.1068	.1132	.1171	.1184	.1171	.1133	.1074	.0997	.0906	.0808	33
18		.0880	.0976	.1057	.1118	.1156	.1169	.1156	.1120	.1062	.0987	32
19		.0666	.0774	.0877	.0970	.1048	.1107	.1144	.1156	.1144	.1109	31
20		.0463	.0564	.0670	.0775	.0875	.0965	.1041	.1098	.1134	.1146	30
21		.0297	.0379	.0471	.0570	.0673	.0776	.0874	.0962	.1035	.1091	29
22		.0176	.0235	.0306	.0387	.0478	.0575	.0676	.0777	.0873	.0959	28
23		.0096	.0135	.0183	.0243	.0313	.0394	.0484	.0580	.0679	.0778	27
24		.0049	.0071	.0102	.0141	.0193	.0249	.0319	.0400	.0489	.0584	26
25		.0023	.0035	.0052	.0075	.0105	.0146	.0195	.0255	.0325	.0405	25
26		.0010	.0016	.0025	.0037	.0055	.0079	.0110	.0150	.0200	.0259	24
27		.0004	.0007	.0011	.0017	.0025	.0039	.0058	.0082	.0113	.0154	23
28		.0001	.0003	.0004	.0007	.0012	.0018	.0028	.0041	.0060	.0084	22
29		.0000	.0001	.0002	.0003	.0005	.0008	.0012	.0019	.0029	.0043	21
30		.0000	.0000	.0001	.0001	.0002	.0003	.0005	.0008	.0013	.0020	20
31		.0000	.0000	.0000	.0000	.0001	.0001	.0002	.0003	.0005	.0009	19
32		.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0001	.0002	.0003	18
33		.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0001	17
		.69	.68	.67	.66	.65	.64	.63	.62	.61	.60	P R

R	P	.41	.42	.43	.44	.45	.46	.47	.48	.49	.50	
8		.0001	.0001	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	42
9		.0003	.0002	.0001	.0001	.0000	.0000	.0000	.0000	.0000	.0000	41
10		.0009	.0006	.0004	.0002	.0001	.0001	.0001	.0000	.0000	.0000	40
11		.0024	.0016	.0010	.0007	.0004	.0003	.0002	.0001	.0001	.0000	39
12		.0054	.0037	.0026	.0017	.0011	.0007	.0005	.0003	.0002	.0001	38
13		.0109	.0079	.0057	.0040	.0027	.0018	.0012	.0008	.0005	.0003	37
14		.0200	.0152	.0113	.0082	.0059	.0041	.0029	.0019	.0013	.0008	36
15		.0334	.0264	.0204	.0155	.0115	.0085	.0061	.0043	.0030	.0020	35
16		.0508	.0418	.0337	.0267	.0207	.0158	.0118	.0086	.0062	.0044	34
17		.0706	.0605	.0508	.0419	.0339	.0269	.0209	.0159	.0119	.0087	33
18		.0899	.0803	.0703	.0604	.0508	.0420	.0340	.0270	.0210	.0160	32
19		.1053	.0979	.0893	.0799	.0700	.0602	.0507	.0419	.0340	.0270	31
20		.1134	.1099	.1044	.0973	.0888	.0795	.0697	.0600	.0506	.0419	30
21		.1126	.1137	.1126	.1092	.1038	.0967	.0884	.0791	.0695	.0598	29
22		.1031	.1086	.1119	.1131	.1119	.1086	.1033	.0963	.0880	.0788	28
23		.0872	.0957	.1028	.1082	.1115	.1126	.1115	.1082	.1029	.0960	27
24		.0682	.0780	.0872	.0956	.1026	.1079	.1112	.1124	.1112	.1080	26
25		.0493	.0587	.0684	.0781	.0873	.0956	.1026	.1079	.1112	.1123	25
26		.0329	.0409	.0497	.0590	.0687	.0783	.0875	.0957	.1027	.1080	24
27		.0203	.0263	.0333	.0412	.0500	.0593	.0690	.0786	.0877	.0960	23
28		.0116	.0157	.0206	.0266	.0336	.0415	.0502	.0596	.0692	.0788	22
29		.0061	.0086	.0118	.0159	.0208	.0268	.0338	.0417	.0504	.0598	21

Table B-8 Binomial distribution (Continued)

30		.0030	.0044	.0062	.0087	.0119	.0160	.0210	.0270	.0339	.0419		20
31		.0013	.0020	.0030	.0044	.0063	.0088	.0120	.0161	.0210	.0270		19
32		.0006	.0009	.0014	.0021	.0031	.0044	.0063	.0088	.0120	.0160		18
33		.0002	.0003	.0006	.0009	.0014	.0021	.0031	.0044	.0063	.0087		17
34		.0001	.0001	.0002	.0003	.0005	.0009	.0014	.0020	.0030	.0044		16
35		.0000	.0000	.0001	.0001	.0002	.0003	.0005	.0009	.0013	.0020		15
36		.0000	.0000	.0000	.0000	.0001	.0001	.0002	.0003	.0005	.0008		14
37		.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0001	.0002	.0003		13
38		.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0001		12
		.59	.58	.57	.56	.55	.54	.53	.52	.51	.50	P	R

N = 100

R	P	.01	.02	.03	.04	.05	.06	.07	.08	.09	.10		
0		.3660	.1326	.0476	.0169	.0059	.0021	.0007	.0002	.0001	.0000		100
1		.3697	.2707	.1471	.0703	.0312	.0131	.0053	.0021	.0008	.0003		99
2		.1849	.2734	.2252	.1450	.0812	.0414	.0198	.0090	.0039	.0016		98
3		.0610	.1823	.2275	.1973	.1396	.0864	.0486	.0254	.0125	.0059		97
4		.0149	.0902	.1706	.1994	.1781	.1338	.0888	.0536	.0301	.0159		96
5		.0029	.0353	.1013	.1595	.1800	.1639	.1283	.0895	.0571	.0339		95
6		.0005	.0114	.0496	.1052	.1500	.1657	.1529	.1233	.0895	.0596		94
7		.0001	.0031	.0206	.0589	.1060	.1420	.1545	.1440	.1188	.0889		93
8		.0000	.0007	.0074	.0285	.0649	.1054	.1352	.1455	.1366	.1148		92
9		.0000	.0002	.0023	.0121	.0349	.0687	.1040	.1293	.1381	.1304		91
10		.0000	.0000	.0007	.0046	.0167	.0399	.0712	.1024	.1243	.1319		90
11		.0000	.0000	.0002	.0016	.0072	.0209	.0439	.0728	.1006	.1199		89
12		.0000	.0000	.0000	.0005	.0028	.0099	.0245	.0470	.0738	.0988		88
13		.0000	.0000	.0000	.0001	.0010	.0043	.0125	.0276	.0494	.0743		87
14		.0000	.0000	.0000	.0000	.0003	.0017	.0058	.0149	.0304	.0513		86
15		.0000	.0000	.0000	.0000	.0001	.0006	.0025	.0074	.0172	.0327		85
16		.0000	.0000	.0000	.0000	.0000	.0002	.0010	.0034	.0090	.0193		84
17		0	.0000	.0000	.0000	.0000	.0001	.0004	.0015	.0044	.0106		83
18		0	.0000	.0000	.0000	.0000	.0000	.0001	.0006	.0020	.0054		82
19		0	.0000	.0000	.0000	.0000	.0000	.0000	.0002	.0009	.0026		81
20		0	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0003	.0012		80
21		0	0	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0005		79
22		0	0	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0002		78
23		0	0	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001		77
		.99	.98	.97	.96	.95	.94	.93	.92	.91	.90	P	R

R	P	.11	.12	.13	.14	.15	.16	.17	.18	.19	.20		
1		.0001	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000		99
2		.0007	.0003	.0001	.0000	.0000	.0000	.0000	.0000	.0000	.0000		98
3		.0027	.0012	.0005	.0002	.0001	.0000	.0000	.0000	.0000	.0000		97
4		.0080	.0038	.0018	.0008	.0003	.0001	.0001	.0000	.0000	.0000		96
5		.0189	.0100	.0050	.0024	.0011	.0005	.0002	.0001	.0000	.0000		95
6		.0369	.0215	.0119	.0063	.0031	.0015	.0007	.0003	.0001	.0001		94
7		.0613	.0394	.0238	.0137	.0075	.0039	.0020	.0009	.0004	.0002		93
8		.0881	.0625	.0414	.0259	.0153	.0086	.0047	.0024	.0012	.0006		92
9		.1112	.0871	.0632	.0430	.0276	.0168	.0098	.0054	.0029	.0015		91
10		.1251	.1080	.0860	.0637	.0444	.0292	.0182	.0108	.0062	.0034		90
11		.1265	.1205	.1051	.0849	.0640	.0454	.0305	.0194	.0118	.0069		89
12		.1160	.1219	.1165	.1025	.0838	.0642	.0463	.0316	.0206	.0128		88
13		.0970	.1125	.1179	.1130	.1001	.0827	.0642	.0470	.0327	.0216		87
14		.0745	.0954	.1094	.1143	.1098	.0979	.0817	.0641	.0476	.0335		86
15		.0528	.0745	.0938	.1067	.1111	.1070	.0960	.0807	.0640	.0481		85
16		.0347	.0540	.0744	.0922	.1041	.1082	.1044	.0941	.0798	.0638		84
17		.0212	.0364	.0549	.0742	.0908	.1019	.1057	.1021	.0924	.0789		83
18		.0121	.0229	.0379	.0557	.0739	.0895	.0998	.1033	.1000	.0909		82
19		.0064	.0135	.0244	.0391	.0563	.0736	.0882	.0979	.1012	.0981		81
20		.0032	.0074	.0148	.0258	.0412	.0567	.0732	.0870	.0962	.0993		80
21		.0015	.0039	.0084	.0160	.0270	.0412	.0571	.0728	.0859	.0946		79
22		.0007	.0019	.0045	.0094	.0171	.0282	.0420	.0574	.0724	.0849		78
23		.0003	.0009	.0023	.0052	.0103	.0182	.0292	.0427	.0576	.0720		77
24		.0001	.0004	.0011	.0027	.0058	.0111	.0192	.0301	.0433	.0577		76
25		.0000	.0002	.0005	.0013	.0031	.0064	.0119	.0201	.0309	.0439		75
26		.0000	.0001	.0002	.0006	.0016	.0035	.0071	.0127	.0209	.0316		74

Table B-8 Binomial distribution (Continued)

27		.0000	.0000	.0001	.0003	.0008	.0018	.0040	.0076	.0134	.0217		73
28		.0000	.0000	.0000	.0001	.0004	.0009	.0021	.0044	.0082	.0141		72
29		.0000	.0000	.0000	.0000	.0002	.0004	.0011	.0024	.0048	.0088		71
30		.0000	.0000	.0000	.0000	.0001	.0002	.0005	.0012	.0027	.0052		70
31		.0000	.0000	.0000	.0000	.0000	.0001	.0002	.0006	.0014	.0029		69
32		.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0003	.0007	.0016		68
33		.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0003	.0008		67
34		.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0002	.0004		66
35		.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0002		65
36		.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001		64
		.89	.88	.87	.86	.85	.84	.83	.82	.81	.80	P	R

R	P	.21	.22	.23	.24	.25	.26	.27	.28	.29	.30		
7		.0001	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000		93
8		.0003	.0001	.0001	.0000	.0000	.0000	.0000	.0000	.0000	.0000		92
9		.0007	.0003	.0002	.0001	.0000	.0000	.0000	.0000	.0000	.0000		91
10		.0018	.0009	.0004	.0002	.0001	.0000	.0000	.0000	.0000	.0000		90
11		.0038	.0021	.0011	.0005	.0003	.0001	.0001	.0000	.0000	.0000		89
12		.0076	.0043	.0024	.0012	.0006	.0003	.0001	.0001	.0000	.0000		88
13		.0136	.0082	.0048	.0027	.0014	.0007	.0004	.0002	.0001	.0000		87
14		.0225	.0144	.0089	.0052	.0030	.0016	.0009	.0004	.0002	.0001		86
15		.0343	.0233	.0152	.0095	.0057	.0033	.0018	.0010	.0005	.0002		85
16		.0484	.0350	.0241	.0159	.0100	.0061	.0035	.0020	.0011	.0006		84
17		.0636	.0487	.0356	.0248	.0165	.0106	.0065	.0038	.0022	.0012		83
18		.0780	.0634	.0498	.0361	.0254	.0171	.0111	.0069	.0041	.0024		82
19		.0895	.0772	.0631	.0492	.0365	.0259	.0177	.0115	.0072	.0044		81
20		.0963	.0881	.0764	.0629	.0493	.0369	.0264	.0182	.0120	.0076		80
21		.0975	.0947	.0869	.0756	.0626	.0494	.0373	.0269	.0186	.0124		79
22		.0931	.0959	.0932	.0858	.0749	.0623	.0495	.0376	.0273	.0190		78
23		.0839	.0917	.0944	.0919	.0847	.0743	.0621	.0495	.0378	.0277		77
24		.0716	.0830	.0905	.0931	.0906	.0837	.0736	.0618	.0496	.0380		76
25		.0578	.0712	.0822	.0893	.0918	.0894	.0828	.0731	.0615	.0496		75
26		.0444	.0579	.0708	.0814	.0883	.0906	.0883	.0819	.0725	.0613		74
27		.0323	.0448	.0580	.0704	.0805	.0873	.0896	.0873	.0812	.0720		73
28		.0224	.0329	.0451	.0580	.0701	.0799	.0864	.0886	.0864	.0804		72
29		.0148	.0231	.0335	.0455	.0580	.0697	.0793	.0855	.0676	.0856		71
30		.0093	.0154	.0237	.0340	.0458	.0580	.0694	.0787	.0847	.0868		70
31		.0056	.0098	.0160	.0242	.0344	.0460	.0580	.0691	.0781	.0840		69
32		.0032	.0060	.0103	.0165	.0248	.0349	.0462	.0579	.0688	.0776		68
33		.0018	.0035	.0063	.0107	.0170	.0252	.0352	.0464	.0579	.0685		67
34		.0009	.0019	.0037	.0067	.0112	.0175	.0257	.0356	.0466	.0579		66
35		.0005	.0010	.0021	.0040	.0070	.0116	.0179	.0261	.0359	.0468		65
36		.0002	.0005	.0011	.0023	.0042	.0073	.0120	.0183	.0265	.0362		64
37		.0001	.0003	.0006	.0012	.0024	.0045	.0077	.0123	.0187	.0268		63
38		.0000	.0001	.0003	.0006	.0013	.0026	.0047	.0079	.0127	.0191		62
39		.0000	.0001	.0001	.0003	.0007	.0015	.0028	.0049	.0082	.0130		61
40		.0000	.0000	.0001	.0002	.0004	.0008	.0016	.0029	.0051	.0085		60
41		.0000	.0000	.0000	.0001	.0002	.0004	.0008	.0017	.0031	.0053		59
42		.0000	.0000	.0000	.0000	.0001	.0002	.0004	.0009	.0018	.0032		58
43		.0000	.0000	.0000	.0000	.0000	.0001	.0002	.0005	.0010	.0019		57
44		.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0002	.0005	.0010		56
45		.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0003	.0005		55
46		.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0001	.0003		54
47		.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0001		53
48		.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001		52
		.79	.78	.77	.76	.75	.74	.73	.72	.71	.70	P	R

R	P	.31	.32	.33	.34	.35	.36	.37	.38	.39	.40		
15		.0001	.0001	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000		85
16		.0003	.0001	.0001	.0000	.0000	.0000	.0000	.0000	.0000	.0000		84
17		.0006	.0003	.0002	.0001	.0000	.0000	.0000	.0000	.0000	.0000		83
18		.0013	.0007	.0004	.0002	.0001	.0000	.0000	.0000	.0000	.0000		82
19		.0025	.0014	.0008	.0004	.0002	.0001	.0000	.0000	.0000	.0000		81
20		.0046	.0027	.0015	.0008	.0004	.0002	.0001	.0001	.0000	.0000		80
21		.0079	.0049	.0029	.0016	.0009	.0005	.0002	.0001	.0001	.0000		79
22		.0127	.0082	.0051	.0030	.0017	.0010	.0005	.0003	.0001	.0001		78
23		.0194	.0131	.0085	.0053	.0032	.0018	.0010	.0006	.0003	.0001		77

Table B-8 Binomial distribution (Continued)

24	.0280	.0198	.0134	.0088	.0055	.0033	.0019	.0011	.0006	.0003	76
25	.0382	.0283	.0201	.0137	.0090	.0057	.0035	.0020	.0012	.0006	75
26	.0496	.0384	.0286	.0204	.0140	.0092	.0059	.0036	.0021	.0012	74
27	.0610	.0495	.0386	.0288	.0207	.0143	.0095	.0060	.0037	.0022	73
28	.0715	.0608	.0495	.0387	.0290	.0209	.0145	.0097	.0062	.0038	72
29	.0797	.0710	.0605	.0495	.0388	.0292	.0211	.0147	.0098	.0063	71
30	.0848	.0791	.0706	.0603	.0494	.0389	.0294	.0213	.0149	.0100	70
31	.0860	.0840	.0785	.0702	.0601	.0494	.0389	.0295	.0215	.0151	69
32	.0833	.0853	.0834	.0779	.0698	.0599	.0493	.0390	.0296	.0217	68
33	.0771	.0827	.0846	.0827	.0774	.0694	.0597	.0493	.0390	.0297	67
34	.0683	.0767	.0821	.0840	.0821	.0769	.0691	.0595	.0492	.0391	66
35	.0578	.0680	.0763	.0816	.0834	.0816	.0765	.0688	.0593	.0491	65
36	.0469	.0578	.0678	.0759	.0811	.0829	.0811	.0761	.0685	.0591	64
37	.0365	.0471	.0578	.0676	.0755	.0806	.0824	.0807	.0757	.0682	63
38	.0272	.0367	.0472	.0577	.0674	.0752	.0802	.0820	.0803	.0754	62
39	.0194	.0275	.0369	.0473	.0577	.0672	.0749	.0799	.0816	.0799	61
40	.0133	.0197	.0277	.0372	.0474	.0577	.0671	.0746	.0795	.0812	60
41	.0087	.0136	.0200	.0280	.0373	.0475	.0577	.0670	.0744	.0792	59
42	.0055	.0090	.0138	.0203	.0282	.0375	.0476	.0576	.0668	.0742	58
43	.0033	.0057	.0092	.0141	.0205	.0285	.0377	.0477	.0576	.0667	57
44	.0019	.0035	.0059	.0094	.0143	.0207	.0287	.0378	.0477	.0576	56
45	.0011	.0020	.0036	.0060	.0095	.0145	.0210	.0289	.0380	.0478	55
46	.0006	.0011	.0021	.0037	.0062	.0098	.0147	.0212	.0290	.0381	54
47	.0003	.0006	.0012	.0022	.0038	.0063	.0099	.0149	.0213	.0292	53
48	.0001	.0003	.0007	.0012	.0023	.0039	.0064	.0101	.0151	.0215	52
49	.0001	.0002	.0003	.0007	.0013	.0023	.0040	.0066	.0102	.0152	51
50	.0000	.0001	.0002	.0004	.0007	.0013	.0024	.0041	.0067	.0103	50
51	.0000	.0000	.0001	.0002	.0004	.0007	.0014	.0025	.0042	.0068	49
52	.0000	.0000	.0000	.0001	.0002	.0004	.0008	.0014	.0025	.0042	48
53	.0000	.0000	.0000	.0000	.0001	.0002	.0004	.0008	.0015	.0026	47
54	.0000	.0000	.0000	.0000	.0000	.0001	.0002	.0004	.0008	.0015	46
55	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0002	.0004	.0008	45
56	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0002	.0004	44
57	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0001	.0002	43
58	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0001	42
59	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	41

Index

Index

Activities, 163
Adverse minimum, 69
Aggregate production planning, 330
Air pollution, 251
Assignment algorithm, 198
Attribute sampling, 435
Autocorrelation, 320
Average outgoing quality, 442

Backorders, 330
Bank auditing, 449
Batch processing, 464
Bayesian quality control, 450
Benefit/cost:
  distribution, 43
  material handling, 100
  for replacement decisions, 76
Biochemical oxygen demand, 250
Branch and bound:
  assignment algorithm, 198
  travelling salesman, 402
Breakeven analysis, 95
Breakeven chart, 96

Capacity decisions, 230
Capital budgeting process, 22
Capital recovery factors, 69
Certainty, decision making under, 10
Challenger, 59
Coefficient of determination, 309
Coefficient of multiple determination, 316
Computers, 463
Construction costs, 223
Continuous process, 5, 92
Control charts:
  attributes, 448
  variables, 445
Correlation, 310
  between investments, 48
Cost of capital, 29
Coterminated plant models, 60
Court scheduling, 425
CPM (*see* Critical path method)
CRAFT, 213
Criterion, 9
Critical path, 166
Critical path method, 175
  linear programming model, 179

Decision variable, 35
Decisions, structure of, 8
Defender, 60
Degradable wastes, 249
Deterioration, 65
Dimensional analysis, 227
Discount factor, 26
Dissolved oxygen, 250
Disturbance term, 305
Double sampling, 442
Dummy activities, 165
Dummy variables, 322

Econometric forecasting, 301
Economic life, 62, 69
Engineering design, 143
Environmental states, 9
EOQ model, 367
Equipment selection, 95
Events, 164
Exponential smoothing, 271

F test, 318
Facilities location, 221
Fixed and variable cost model, 239
Fixed order inventory system, 385
Flow process chart, 99
Flow shop sequencing, 398
Futurity, 62

Gantt chart, 393
Graphical linear programming, 480

Health delivery, 93
Heuristic:
  definition, 16
  facilities layout, 213
  line balancing, 121
  production planning, 353
  project scheduling with scarce resources, 182
Holt, Modigliani, and Simon model, 346
Homoscedasticity, 320
Human relations, 143

Information systems, 459
Integer programming, 497
Intermittent process, 5, 92
Internal rate of return, 28
Inventory, 333, 362
Investment, 21

Job enlargement, 144
Job enrichment, 151
Job participation, 148
Job shop scheduling and control, 419
Job shop sequencing, 416
Judgemental forecasting, 272

Labor standards, 117, 130
Lagged variables, 325
Lateness in sequencing, 398
Lead time, 364
Limited resource scheduling, 181
Line of least squares, 302
Line balancing, 114
Linear programming:
  aggregate production planning, 339
  capital budgeting, 35
  critical path analysis, 179
  graphical analysis, 484
  limited resource scheduling, 183
  line balancing, 118
  network analysis, 168
  simplex method, 484

Machine loading, 392
Make or buy, 104
Man machine analysis, 126
Management information systems, 459
Material handling, 100
Medical, 466
Micromotion analysis, 125
Motivation, 156
Multicollinearity, 321
Multifacility location problems, 230
Multiple rates of return, 32
Multiprogramming, 464
Multivariate regression, 314
Mutually exclusive investments, 33, 41

Network design, 163, 164
Noise pollution, 252
Nondegradable wastes, 249
Nonlinear production costs, 346
Nonlinear programming, 497
Nonlinear regression, 311
Northwest corner method, 233

Obsolescence, 66
Operating characteristic curve, 436

Operation charts, 125
Operation process chart, 97
Operations management, definition of, 5
Outcome, 9
Overtime, 332
Oxygen sag, 250

Paced line, 115
Parabolic functions, 313
Parameters, 14
Payback, 23
Performance, 156
Periodic review inventory system, 385
PERT, 169
  simulation, 173
Physiological factors, 129
Pollution, 246
Pollution standards, 223
Power, 222
Power function, 312
Precedence, 113
Predetermined time standards, 135
Present value, 25
Pricing, 354
Process control, 444
Process design:
  macro, 91
  micro, 92
Process layout:
  definition, 92
  dependence, 208
  independence, 194
Process R&D, 90
Processing time, 397
Product development, 90
Product final design, 91
Product flow process chart, 99
Product layout, 92
Product planning, 90
Progressive patient care, 93
Project, 5, 162
Promotion, 354

Quality control, 432
Quantity discounts, 372

Random numbers, 499
Rating factor, 132
Receiving, 222
Regulation of pollution, 260
Remote batch, 464
Replications, 502
Risk, decision making under, 10

Sampling risk, 434
Satisfaction, 156
Schematic aids, 96
Scientific management, 143
Seasonal forecasting, 284
Select time, 131
Semipaced line, 115
Sensitivity analysis:
  benefit/cost distribution, 47
  in linear programming, 496
Sequencing, 394
Sequential sampling, 442
Shadow prices, 492
Shipping, 222
Simple exponential smoothing, 276
Simple linear regression, 302
Simplex, 484
Simulation:
  in-process inventories, 115
  introduction to, 498
  investments, 42
  job shop scheduling, 419
  materials handling, 100
  PERT network, 173
Smoothing constants, 277, 289
Standard time, 113, 132
Stochastic programming, 497
Stockouts, 366, 375
Strategies, 9
Subcontracting, 332, 333, 336
Subsidy for pollution, 261
Supervision, 157

Taylor, Frederick, 3
Terborgh, George, 65
Time sharing, 465
Time study, 131
Time value of money, 25
Transformation process, 89
Transportation method, 233
Travelling salesman problem, 398
Two-factor theory, 151

Uncertainty, decision making under, 10
Uniform annual equivalent, 69
Utility, 12

Variability, operation, 115
Variable sampling, 444

Waste disposal, 252
Waste disposal costs, 253
Waste treatment, 258
Water pollution, 249
Weighted average, 274
Welfare damages of pollution, 256
Work flow, 100
Work methods, 117, 125
Work sampling, 135
Work station selection, 95

Zoning, 222

OPERATIONS
MANAGEMENT